MTH5112 Linear Algebra I, 2009–2010

Oscar Bandtlow

30 March 2010

Contents

Chapter 1

Systems of Linear Equations

Systems of linear equations are the bread and butter of linear algebra. They arise in many areas of the sciences, including physics, engineering, business, economics, and sociology. Their systematic study also provided part of the motivation for the development of modern linear algebra at the end of the 19th century.

The material in this chapter will be familiar from Geometry I, where systems of linear equations have already been discussed in some detail. As this chapter is fundamental for what is to follow, you want to make sure that the basic ideas are hardwired in your brain for the rest of term!

1.1 Basic terminology and examples

A **linear equation in n unknowns** is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b\,,$$

where $a_1, \ldots, a_n$ and b are given real numbers and $x_1, \ldots, x_n$ are variables.

A **system of m linear equations in n unknowns** is a collection of equations of the form

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

where the a_{ij}'s and b_i's are all real numbers. We also call such systems $m \times n$ **systems**.

Example 1.1.

$$(a)\ \begin{array}{rcrcl} 2x_1 & + & x_2 & = & 4 \\ 3x_1 & + & 2x_2 & = & 7 \end{array} \qquad (b)\ \begin{array}{rcrcrcl} x_1 & + & x_2 & - & x_3 & = & 3 \\ 2x_1 & - & x_2 & + & x_3 & = & 6 \end{array} \qquad (c)\ \begin{array}{rcrcl} x_1 & - & x_2 & = & 0 \\ x_1 & + & x_2 & = & 3\,. \\ & & x_2 & = & 1 \end{array}$$

(a) is a 2×2 system, (b) is a 2×3 system, and (c) is a 3×2 system.

A **solution** of an $m \times n$ system is an ordered n-tuple $(x_1, x_2, \ldots, x_n)$ that satisfies all equations of the system.

Example 1.2. $(1, 2)$ is a solution of Example 1.1 (a).
For each $\alpha \in \mathbb{R}$, the 3-tuple $(3, \alpha, \alpha)$ is a solution of Example 1.1 (b) (CHECK!).
Example 1.1 (c) has no solution, since, on the one hand $x_2 = 1$ by the last equation, but the first equation implies $x_1 = 1$, while the second equation implies $x_1 = 2$, which is impossible.

A system with no solution is called **inconsistent**, while a system with at least one solution is called **consistent**.

The set of all solutions of a system is called its **solution set**, which may be empty if the system is inconsistent.

The basic problem we want to address in this section is the following: given an arbitrary $m \times n$ system, determine its solution set. Later on, we will discuss a procedure that provides a complete and practical solution to this problem (the so-called 'Gaussian algorithm'). Before we encounter this procedure, we require a bit more terminology.

Definition 1.3. Two $m \times n$ systems are said to be **equivalent**, if they have the same solution set.

Example 1.4. Consider the two systems

$$(a)\quad \begin{array}{rcrcrcl} 5x_1 & - & x_2 & + & 2x_3 & = & -3 \\ & & x_2 & & & = & 2 \\ & & & & 3x_3 & = & 6 \end{array} \qquad (b)\quad \begin{array}{rcrcrcl} 5x_1 & - & x_2 & + & 2x_3 & = & -3 \\ -5x_1 & + & 2x_2 & - & 2x_3 & = & 5 \\ 5x_1 & - & x_2 & + & 5x_3 & = & 3 \end{array}.$$

(a) is easy to solve: looking at the last equation we find first that $x_3 = 2$; the second from the bottom implies $x_2 = 2$; and finally the first one yields $x_1 = (-3 + x_2 - 2x_3)/5 = -1$. So the solution set of this system is $\{(-1, 2, 2)\}$.

To find the solution of (b), add the first and the second equation. Then $x_2 = 2$, while subtracting the first from the third equation gives $3x_3 = 6$, that is $x_3 = 2$. Finally, the first equation now gives $x_1 = (-3 + x_2 - 2x_3)/5 = -1$, so the solution set is again $\{(-1, 2, 2)\}$.

Thus the systems (a) and (b) are equivalent.

In solving system (b) above we have implicitly used the following important observation:

Lemma 1.5. *The following operations do not change the solution set of a linear system:*

(i) interchanging two equations;

(ii) multiplying an equation by a non-zero scalar;

(iii) adding a multiple of one equation to another.

Proof. (i) and (ii) are obvious. The proof of (iii) was an exercise in Geometry I (Coursework 4, Exercise 3) for the special case of two equations in three unknowns. The general case can be proved in exactly the same way and is left as an exercise. □

We shall see shortly how to use the above operations systematically to obtain the solution set of any given linear system. Before doing so, however, we introduce a useful short-hand.

Given an $m \times n$ linear system

$$\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ \vdots & & \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array}$$

we call the array

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right)$$

the **augmented matrix** of the linear system, and the $m \times n$ matrix

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

the **coefficient matrix** of the linear system.

Example 1.6.

$$\text{system: } \begin{array}{rcrcrcl} 3x_1 & + & 2x_2 & - & x_3 & = & 5 \\ 2x_1 & & & + & x_3 & = & -1 \end{array} \qquad \text{augmented matrix: } \left(\begin{array}{ccc|c} 3 & 2 & -1 & 5 \\ 2 & 0 & 1 & -1 \end{array}\right).$$

A system can be solved by performing operations on the augmented matrix. Corresponding to the three operations given in Lemma 1.5 we have the following three operations that can be applied to the augmented matrix, called **elementary row operations**.

Definition 1.7 (Elementary row operations)**.**

Type I interchanging two rows;
Type II multiplying a row by a non-zero scalar;
Type III adding a multiple of one row to another row.

1.2 Gaussian elimination

Gaussian elimination is a systematic procedure to determine the solution set of a given linear systen. The basic idea is to perform elementary row operations on the corresponding augmented matrix bringing it to a simpler form from which the solution set is readily obtained.

The simple form alluded to above is given in the following definition.

Definition 1.8. A matrix is said to be in **row echelon form** if it satisfies the following three conditions:

(i) All zero rows (consisting entirely of zeros) are at the bottom.

(ii) The first non-zero entry from the left in each nonzero row is a 1, called the **leading** 1 for that row.

(iii) Each leading 1 is to the right of all leading 1's in the rows above it.

A row echelon matrix is said to be in **reduced row echelon form** if, in addition it satisfies the following condition:

(iv) Each leading 1 is the only nonzero entry in its column

Roughly speaking, a matrix is in row echelon form if the leading 1's form an echelon (that is, a 'steplike') pattern.

Example 1.9. Matrices in row echelon form:

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

Matrices in reduced row echelon form:

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

The variables corresponding to the leading 1's of the augmented matrix in row echelon form will be referred to as the **leading variables**, the remaining ones as the **free variables**.

Example 1.10.

(a) $\left(\begin{array}{cccc|c} 1 & 2 & 3 & -4 & 6 \\ 0 & 0 & 1 & 2 & 3 \end{array}\right)$.
Leading variables: x_1 and x_3; free variables: x_2 and x_4.

(b) $\left(\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & 3 \end{array}\right)$.
Leading variables: x_1 and x_2; no free variables.

Note that if the augmented matrix of a system is in row echelon form, the solution set is easily obtained.

Example 1.11. Determine the solution set of the systems given by the following augmented matrices in row echelon form:

$$\text{(a)}\ \left(\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{array}\right), \quad \text{(b)}\ \left(\begin{array}{cccc|c} 1 & -2 & 0 & 1 & 2 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Solution. (a) The corresponding system is

$$\begin{array}{rcrcl} x_1 & + & 3x_2 & = & 2 \\ & & 0 & = & 1 \end{array}$$

so the system is inconsistent and the solution set is empty.

(b) The corresponding system is

$$\begin{array}{rcrcrcl} x_1 & - & 2x_2 & + & x_4 & = & 2 \\ & & x_3 & - & 2x_4 & = & 1 \\ & & & & 0 & = & 0 \end{array}$$

We can express the leading variables in terms of the free variables x_2 and x_4. So set $x_2 = \alpha$ and $x_4 = \beta$, where α and β are arbitrary real numbers. The second line now tells us that $x_3 = 1 + 2x_4 = 1 + 2\beta$, and then the first line that $x_1 = 2 + 2x_2 - x_4 = 2 + 2\alpha - \beta$. Thus the solution set is $\{\, (2 + 2\alpha - \beta, \alpha, 1 + 2\beta, \beta) \mid \alpha, \beta \in \mathbb{R} \,\}$. □

It turns out that every matrix can be brought into row echelon form using only elementary row operations. The procedure is known as the

Gaussian algorithm:

Step 1 If the matrix consists entirely of zeros, stop — it is already in row echelon form.

Step 2 Otherwise, find the the first column from the left containing a non-zero entry (call it a), and move the row containing that entry to the top position.

Step 3 Now multiply that row by $1/a$ to create a leading 1.

Step 4 By subtracting multiples of that row from rows below it, make each entry below the leading 1 zero.

This completes the first row. All further operations are carried out on the other rows.

Step 5 Repeat steps 1-4 on the matrix consisting of the remaining rows

The process stops when either no rows remain at Step 5 or the remaining rows consist of zeros.

Example 1.12. Solve the following system using the Gaussian algorithm:

$$\begin{array}{rcrcrl} & & x_2 & + & 6x_3 & = 4 \\ 3x_1 & - & 3x_2 & + & 9x_3 & = -3 \\ 2x_1 & + & 2x_2 & + & 18x_3 & = 8 \end{array}$$

Solution. Performing the Gaussian algorithm on the augmented matrix gives:

$$\left(\begin{array}{ccc|c} 0 & 1 & 6 & 4 \\ 3 & -3 & 9 & -3 \\ 2 & 2 & 18 & 8 \end{array}\right) \sim \begin{matrix} \\ R_1 \leftrightarrow R_2 \end{matrix} \left(\begin{array}{ccc|c} 3 & -3 & 9 & -3 \\ 0 & 1 & 6 & 4 \\ 2 & 2 & 18 & 8 \end{array}\right) \sim \begin{matrix} \\ \frac{1}{3}R_1 \end{matrix} \left(\begin{array}{ccc|c} 1 & -1 & 3 & -1 \\ 0 & 1 & 6 & 4 \\ 2 & 2 & 18 & 8 \end{array}\right)$$

$$\sim \begin{matrix} \\ R_3 - 2R_1 \end{matrix} \left(\begin{array}{ccc|c} 1 & -1 & 3 & -1 \\ 0 & 1 & 6 & 4 \\ 0 & 4 & 12 & 10 \end{array}\right) \sim \begin{matrix} \\ R_3 - 4R_2 \end{matrix} \left(\begin{array}{ccc|c} 1 & -1 & 3 & -1 \\ 0 & 1 & 6 & 4 \\ 0 & 0 & -12 & -6 \end{array}\right) \sim \begin{matrix} \\ -\frac{1}{12}R_3 \end{matrix} \left(\begin{array}{ccc|c} 1 & -1 & 3 & -1 \\ 0 & 1 & 6 & 4 \\ 0 & 0 & 1 & \frac{1}{2} \end{array}\right),$$

where the last matrix is now in row echelon form. The corresponding system reads:

$$\begin{array}{rcrcrl} x_1 & - & x_2 & + & 3x_3 & = -1 \\ & & x_2 & + & 6x_3 & = 4 \\ & & & & x_3 & = \frac{1}{2} \end{array}$$

Leading variables are x_1, x_2 and x_3; there are no free variables. The last equation now implies $x_3 = \frac{1}{2}$; the second equation from bottom yields $x_2 = 4 - 6x_3 = 1$ and finally the first equation yields $x_1 = -1 + x_2 - 3x_3 = -\frac{3}{2}$. Thus the solution is $\{(-\frac{3}{2}, 1, \frac{1}{2})\}$. □

A variant of the Gauss algorithm is the Gauss-Jordan algorithm, which brings a matrix to reduced row echelon form:

Gauss-Jordan algorithm

Step 1 Bring matrix to row echelon form using the Gaussian algorithm.

Step 2 Find the row containing the first leading 1 from the right, and add suitable multiples of this row to the rows above it to make each entry above the leading 1 zero.

This completes the first non-zero row from the bottom. All further operations are carried out on the rows above it.

Step 3 Repeat steps 1-2 on the matrix consisting of the remaining rows.

Example 1.13. Solve the following system using the Gauss-Jordan algorithm:

$$\begin{array}{rcrcrcrcrcl} x_1 & + & x_2 & + & x_3 & + & x_4 & + & x_5 & = & 4 \\ x_1 & + & x_2 & + & x_3 & + & 2x_4 & + & 2x_5 & = & 5 \\ x_1 & + & x_2 & + & x_3 & + & 2x_4 & + & 3x_5 & = & 7 \end{array}$$

Solution. Performing the Gauss-Jordan algorithm on the augmented matrix gives:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 4 \\ 1 & 1 & 1 & 2 & 2 & 5 \\ 1 & 1 & 1 & 2 & 3 & 7 \end{array}\right) \begin{array}{r} \\ \sim R_2 - R_1 \\ R_3 - R_1 \end{array} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 4 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 2 & 3 \end{array}\right) \sim \begin{array}{r} \\ \\ R_3 - R_2 \end{array} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 4 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right)$$

$$\begin{array}{r} R_1 - R_3 \\ \sim R_2 - R_3 \\ \\ \end{array} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right) \sim \begin{array}{r} R_1 - R_2 \\ \\ \\ \end{array} \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right),$$

where the last matrix is now in reduced row echelon form. The corresponding system reads:

$$\begin{array}{rcrcrcl} x_1 & + & x_2 & + & x_3 & = & 3 \\ & & & & x_4 & = & -1 \\ & & & & x_5 & = & 2 \end{array}$$

Leading variables are x_1, x_4, and x_5; free variables x_2 and x_3. Now set $x_2 = \alpha$ and $x_3 = \beta$, and solve for the leading variables starting from the last equation. This yields $x_5 = 2$, $x_4 = -1$, and finally $x_1 = 3 - x_2 - x_3 = 3 - \alpha - \beta$. Thus the solution set is $\{\, (3 - \alpha - \beta, \alpha, \beta, -1, 2) \mid \alpha, \beta \in \mathbb{R} \,\}$. □

We have just seen that any matrix can be brought to (reduced) row echelon form using only elementary row operations, and moreover that there is an explicit procedure to achieve this (namely the Gaussian and Gauss-Jordan algorithm). We record this important insight for later use.

Theorem 1.14.

(a) Every matrix can be brought to row echelon form by a series of elementary row operations.

(b) Every matrix can be brought to reduced row echelon form by a series of elementary row operations.

Proof. For (a):apply the Gaussian algorithm; for (b): apply the Gauss-Jordan algorithm. □

Remark 1.15. It can be shown (but not in this module) that the reduced row echelon form of a matrix is unique. A moment's thought (to which you are warmly invited) shows that this is not the case for the row echelon form.

The remark above implies that if a matrix is brought to reduced row echelon form by any sequence of elementary row operations (that is, not necessarily by those prescribed by the Gauss-Jordan algorithm) the leading ones will nevertheless always appear in the same positions. As a consequence, the following definition makes sense.

Definition 1.16. A **pivot position** in a matrix A is a location that corresponds to a leading 1 in the reduced row echelon form of A. A **pivot column** is a column of A that contains a pivot position.

Example 1.17. Let

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 4 \\ 1 & 1 & 1 & 2 & 2 & 5 \\ 1 & 1 & 1 & 2 & 3 & 7 \end{pmatrix}.$$

By Example 1.13 the reduced row echelon form of A is

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

Thus the pivot positions of A are the $(1,1)$-entry, the $(2,4)$-entry, and the $(3,5)$-entry and the pivot columns of A are columns 1, 4, and 5.

The notion of a pivot position and a pivot column will come in handy later in the module.

1.3 Special classes of linear systems

In this last section of our first chapter we'll have a look at a number of special types of linear systems and derive the first important consequences of the fact that every matrix can be brought to row echelon form by a series of elementary row operations.

We start with the following classification of linear systems:

Definition 1.18. An $m \times n$ linear system is said to be

- **overdetermined** if it has more equations than unknowns (i.e. $m > n$);
- **underdetermined** if it has fewer equations than unknowns (i.e. $m < n$).

Note that overdetermined systems are usually (but not necessarily) inconsistent. Underdetermined systems may or may not be consistent. However, if they are consistent, then they necessarily have infinitely many solutions:

Theorem 1.19. *If an underdetermined system is consistent, it must have infinitely many solutions.*

Proof. Note that the row echelon form of the augmented matrix of the system has $r \leq m$ non-zero rows. Thus there are r leading variables, and consequently $n - r \geq n - m > 0$ free variables. □

Another useful classification of linear systems is the following:

Definition 1.20. A linear system

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

is said to be **homogeneous** if $b_i = 0$ for all i. Otherwise it is said to be **inhomogeneous**.

Given an inhomogeneous system (1.1), call the system obtained by setting all b_i's to zero, the **associated homogeneous system** .

Example 1.21.

$$\underbrace{\begin{array}{rcrcrcl} 3x_1 & + & 2x_2 & + & 5x_3 & = & 2 \\ 2x_1 & - & x_2 & + & x_3 & = & 5 \end{array}}_{\text{inhomogeneous system}} \qquad \underbrace{\begin{array}{rcrcrcl} 3x_1 & + & 2x_2 & + & 5x_3 & = & 0 \\ 2x_1 & - & x_2 & + & x_3 & = & 0 \end{array}}_{\text{associated homogeneous system}}$$

The first observation about homogeneous systems is that they always have a solution, the so-called **trivial** or **zero** solution: $(0, 0, \ldots, 0)$.

For later use we record the following useful consequence of the previous theorem on consistent homogeneous systems:

Theorem 1.22. *An underdetermined homogeneous system always has non-trivial solutions.*

Proof. We just observed that a homogeneous systems is consistent. Thus, if the system is underdetermined and homogeneous, it must have infinitely many solutions by Theorem 1.19, hence, in particular, it must have a non-zero solution. □

Our final result in this section is devoted to the special case of $n \times n$ systems. For such systems there is a delightful characterisation of the existence and uniqueness of solutions of a given system in terms of the associated homogeneous systems. At the same time, the proof of this result serves as another illustration of the usefulness of the row echelon form for theoretical purposes.

Theorem 1.23. *An $n \times n$ system is consistent and has a unique solution, if and only if the only solution of the associated homogeneous system is the zero solution.*

Proof. Follows from the following two observations:

- The same sequence of elementary row operations that brings the augmented matrix of a system to row echelon form, also brings the augmented matrix of the associated homogeneous system to row echelon form, and vice versa.
- An $n \times n$ system in row echelon form has a unique solution precisely if there are n leading variables.

Thus, if an $n \times n$ system is consistent and has a unique solution, the corresponding homogeneous system must have a unique solution, which is necessarily the zero solution.

Conversely, if the associated homogeneous system of a given system has the zero solution as its unique solution, then the original inhomogeneous system must have a solution, and this solution must be unique. □

Chapter 2

Matrix Algebra

2.1 Revision from Geometry I

Recall that an $m \times n$ **matrix** A is a rectangular array of scalars (real numbers)

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

We write $A = (a_{ij})_{m\times n}$ or simply $A = (a_{ij})$ to denote an $m \times n$ matrix whose (i, j)-entry is a_{ij}, i.e. a_{ij} is the i-th row and in the j-th column.

If $A = (a_{ij})_{m\times n}$ we say that A has **size** $m \times n$. An $n \times n$ matrix is said to be **square**.

Example 2.1. If

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

then A is a matrix of size 2×3. The $(1, 2)$-entry of A is 3 and the $(2, 3)$-entry of A is 0.

Definition 2.2 (Equality)**.** Two matrices A and B are **equal** and we write $A = B$ if they have the same size and $a_{ij} = b_{ij}$ where $A = (a_{ij})$ and $B = (b_{ij})$.

Definition 2.3 (Scalar multiplication)**.** If $A = (a_{ij})_{m\times n}$ and α is a scalar, then αA (the **scalar product of** α **and** A) is the $m \times n$ matrix whose (i, j)-entry is αa_{ij}.

Definition 2.4 (Addition)**.** If $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{m\times n}$ then the **sum** $A + B$ of A and B is the $m \times n$ matrix whose (i, j)-entry is $a_{ij} + b_{ij}$.

Example 2.5. Let

$$A = \begin{pmatrix} 2 & 3 \\ -1 & 2 \\ 4 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 1 \\ 2 & 3 \\ -2 & 1 \end{pmatrix}.$$

Then

$$3A + 2B = \begin{pmatrix} 6 & 9 \\ -3 & 6 \\ 12 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 4 & 6 \\ -4 & 2 \end{pmatrix} = \begin{pmatrix} 6 & 11 \\ 1 & 12 \\ 8 & 2 \end{pmatrix}.$$

Definition 2.6 (Zero matrix)**.** We write $O_{m\times n}$ or simply O (if the size is clear from the context) for the $m \times n$ matrix all of whose entries are zero, and call it a **zero matrix**.

Scalar multiplication and addition of matrices satisfy the following rules proved in Geometry I:

Theorem 2.7. *Let A, B and C be matrices of the same size, and let α and β be scalars. Then:*

(a) $A + B = B + A$;

(b) $A + (B + C) = (A + B) + C$;

(c) $A + O = A$;

(d) $A + (-A) = O$, *where* $-A = (-1)A$;

(e) $\alpha(A + B) = \alpha A + \alpha B$;

(f) $(\alpha + \beta)A = \alpha A + \beta A$;

(g) $(\alpha\beta)A = \alpha(\beta A)$;

(h) $1A = A$.

Example 2.8. Simplify $2(A + 3B) - 3(C + 2B)$, where A, B, and C are matrices with the same size.

Solution.

$$2(A + 3B) - 3(C + 2B) = 2A + 2 \cdot 3B - 3C - 3 \cdot 2B = 2A + 6B - 3C - 6B = 2A - 3C\,.$$

□

Definition 2.9 (Matrix multplication)**.** If $A = (a_{ij})$ is an $m \times n$ matrix and $B = (b_{ij})$ is an $n \times p$ matrix then the **product** AB of A and B is the $m \times p$ matrix $C = (c_{ij})$ with

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

Example 2.10. Compute the $(1, 3)$-entry and the $(2, 4)$-entry of AB, where

$$A = \begin{pmatrix} 3 & -1 & 2 \\ 0 & 1 & 4 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & 1 & 6 & 0 \\ 0 & 2 & 3 & 4 \\ -1 & 0 & 5 & 8 \end{pmatrix}\,.$$

Solution.

$(1, 3)$-entry: $3 \cdot 6 + (-1) \cdot 3 + 2 \cdot 5 = 25$;

$(2, 4)$-entry: $0 \cdot 0 + 1 \cdot 4 + 4 \cdot 8 = 36$. □

Definition 2.11 (Identity matrix)**.** An **identity matrix** I is a square matrix with 1's on the diagonal and zeros elsewhere. If we want to emphasise its size we write I_n for the $n \times n$ identity matrix.

Matrix multiplication satisfies the following rules proved in Geometry I:

Theorem 2.12. *Assume that α is a scalar and that A, B, and C are matrices so that the indicated operations can be performed. Then:*

(a) $IA = A$ *and* $BI = B$;

(b) $A(BC) = (AB)C$;

(c) $A(B + C) = AB + AC$;

(d) $(B + C)A = BA + CA$;

(e) $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

Notation 2.13.

- Since $A(BC) = (AB)C$, we can omit the brackets and simply write ABC and similarly for products of more than three factors.
- If A is a square matrix we write $A^k = \underbrace{AA\cdots A}_{k \text{ factors}}$ for the k-**th power** of A.

Warning: In general $AB \neq BA$, even if AB and BA have the same size!

Example 2.14.

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

but

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Definition 2.15. If A and B are two matrices with $AB = BA$, then A and B are said to **commute**.

Finally we recall the notion of an inverse of a matrix.

Definition 2.16. If A is a square matrix, a matrix B is called an **inverse** of A if

$$AB = I \quad \text{and} \quad BA = I\,.$$

A matrix that has an inverse is called **invertible**.

Note that not every matrix is invertible. For example the matrix

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

cannot have an inverse since for any 2×2 matrix $B = (b_{ij})$ we have

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ 0 & 0 \end{pmatrix} \neq I_2\,.$$

Later on in this chapter we shall discuss an algorithm that lets us decide whether a matrix is invertible and at the same furnishes an inverse if the matrix is invertible.

It turns out that if a matrix is invertible its inverse is uniquely determined:

Theorem 2.17. *If B and C are both inverses of A, then $B = C$.*

Proof. Since B and C are inverses of A we have $AB = I$ and $CA = I$. Thus

$$B = IB = (CA)B = C(AB) = CI = C\,.$$

□

If A is an invertible matrix, the unique inverse of A is denoted by A^{-1}. Hence A^{-1} (if it exists!) is a square matrix of the same size as A with the property that

$$AA^{-1} = A^{-1}A = I\,.$$

Note that the above equality implies that if A is invertible, then its inverse A^{-1} is also invertible with inverse A, that is,

$$(A^{-1})^{-1} = A\,.$$

Slightly deeper is the following result:

Theorem 2.18. *If A and B are invertible matrices of the same size, then AB is invertible and*

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

Proof. Observe that

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

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I\,.$$

Thus, by definition of invertibility, AB is invertible with inverse $B^{-1}A^{-1}$. □

2.2 Transpose of a matrix

The first new concept we encounter is the following:

Definition 2.19. The **transpose** of an $m \times n$ matrix $A = (a_{ij})$ is the $n \times m$ matrix $B = (b_{ij})$ given by

$$b_{ij} = a_{ji}$$

The transpose of A is denoted by A^T.

Example 2.20.

$$\text{(a)}\quad A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

$$\text{(b)}\quad B = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix} \Rightarrow B^T = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}$$

Matrix transposition satisfies the following rules:

Theorem 2.21. *Assume that α is a scalar and that A, B, and C are matrices so that the indicated operations can be performed. Then:*

(a) $(A^T)^T = A$;

(b) $(\alpha A)^T = \alpha(A^T)$;

(c) $(A+B)^T = A^T + B^T$;

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

Proof. (a) is obvious while (b) and (c) are proved as Exercise 6 in Coursework 2. For the proof of (d) assume $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{n\times p}$ and write $A^T = (\tilde{a}_{ij})_{n\times m}$ and $B^T = (\tilde{b}_{ij})_{p\times n}$ where

$$\tilde{a}_{ij} = a_{ji} \quad \text{and} \quad \tilde{b}_{ij} = b_{ji}\,.$$

Notice that $(AB)^T$ and $B^T A^T$ have the same size, so it suffices to show that they have the same entries. Now, the (i,j)-entry of $B^T A^T$ is

$$\sum_{k=1}^{n} \tilde{b}_{ik}\tilde{a}_{kj} = \sum_{k=1}^{n} b_{ki}a_{jk} = \sum_{k=1}^{n} a_{jk}b_{ki}\,,$$

which is the (j,i)-entry of AB, that is, the (i,j)-entry of $(AB)^T$. Thus $B^T A^T = (AB)^T$. □

Transposition ties in nicely with invertibility:

Theorem 2.22. *Let A be invertible. Then A^T is invertible and*

$$(A^T)^{-1} = (A^{-1})^T\,.$$

Proof. See Exercise 8 in Coursework 2. □

2.3 Special types of square matrices

In this section we briefly introduce a number of special classes of matrices which will be studied in more detail later in this course.

Definition 2.23. A matrix is said to be **symmetric** if $A^T = A$.

Note that a symmetric matrix is necessarily square.

Example 2.24.

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

$$\text{not symmetric:} \quad \begin{pmatrix} 2 & 2 & 4 \\ 2 & 2 & 3 \\ 1 & 3 & 5 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Innocent as this definition may seem, symmetric matrices play an important role in many parts of pure and applied Mathematics as well as in some of the 'hard' sciences. Some of the reasons for this will become clearer towards the end of this course, when we shall study symmetric matrices in much more detail.

Some other useful classes of square matrices are the triangular ones, which will also play a role later on in the course.

Definition 2.25. A square matrix $A = (a_{ij})$ is said to be

upper triangular	if $a_{ij} = 0$ for $i > j$;
strictly upper triangular	if $a_{ij} = 0$ for $i \geq j$;
lower triangular	if $a_{ij} = 0$ for $i < j$;
strictly lower triangular	if $a_{ij} = 0$ for $i \leq j$;
diagonal	if $a_{ij} = 0$ for $i \neq j$.

If $A = (a_{ij})$ is a square matrix of size $n \times n$, we call $a_{11}, a_{22}, \ldots, a_{nn}$ the **diagonal entries** of A. So, informally speaking, a matrix is upper triangular if all the entries below the diagonal entries are zero, and it is strictly upper triangular if all entries below the diagonal entries and the diagonal entries itself are zero. Similarly for (strictly) lower triangular matrices.

Example 2.26.

$$\text{upper triangular: } \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}, \quad \text{diagonal: } \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

$$\text{strictly lower triangular: } \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 2 & 3 & 0 \end{pmatrix}.$$

We close this section with the following two observations:

Theorem 2.27. *The sum and product of two upper triangular matrices of the same size is upper triangular.*

Proof. See Exercise 5, Coursework 2. □

2.4 Linear systems in matrix notation

We shall now have another look at systems of linear equations. The added spice in this discussion will be that we now use the language of matrices to study them. More precisely, we shall now introduce two equivalent ways of writing systems of linear equations. Both reformulations will in their own way shed some light on both linear systems and matrices.

Before discussing these reformulations let us recall from Geometry I that an $n \times 1$ matrix

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

is called a **column vector of dimension** n, or simply an n-**vector**. The collection of all n-vectors is denoted by $\mathbb{R}^n$. Thus:

$$\mathbb{R}^n = \left\{ \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \,\middle|\, a_1, a_2, \cdots, a_n \in \mathbb{R} \right\}.$$

Suppose now that we are given an $m \times n$ linear system

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &=& b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &=& b_2 \\ \vdots & & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &=& b_m \end{array}. \tag{2.1}$$

The first reformulation is based on the observation that we can write this system more succinctly as a single matrix equation

$$A\mathbf{x} = \mathbf{b}\,, \tag{2.2}$$

where

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \text{ and } \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m,$$

and where $A\mathbf{x}$ is interpreted as the matrix product of A and $\mathbf{x}$.

Example 2.28. Using matrix notation the system

$$\begin{array}{rcrcrcl} 2x_1 & - & 3x_2 & + & x_3 & = & 2 \\ 3x_1 & & & - & x_3 & = & -1 \end{array}$$

can be written

$$\underbrace{\begin{pmatrix} 2 & -3 & 1 \\ 3 & 0 & -1 \end{pmatrix}}_{=A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{=\mathbf{x}} = \underbrace{\begin{pmatrix} 2 \\ -1 \end{pmatrix}}_{=\mathbf{b}}.$$

Apart from obvious notational economy, writing (2.1) in the form (2.2) has a number of other advantages which will become clearer shortly.

The other useful way of writing (2.1) is the following: with A and $\mathbf{x}$ as before we have

$$A\mathbf{x} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix} = x_1 \underbrace{\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}}_{=\mathbf{a}_1} + \cdots + x_n \underbrace{\begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}}_{=\mathbf{a}_n},$$

where $\mathbf{a}_j \in \mathbb{R}^m$ is the j-th column of A.

Thus the linear system (2.1) can also be represented as a matrix (or vector) equation of the form

$$x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}\,. \tag{2.3}$$

Example 2.29. The linear system

$$\begin{array}{rcrcrcl} 2x_1 & - & 3x_2 & - & 2x_3 & = & 5 \\ 5x_1 & - & 4x_2 & + & 2x_3 & = & 6 \end{array}$$

can also be written as

$$x_1 \begin{pmatrix} 2 \\ 5 \end{pmatrix} + x_2 \begin{pmatrix} 3 \\ -4 \end{pmatrix} + x_3 \begin{pmatrix} -2 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix} . \tag{2.4}$$

Sums such as the left-hand side of (2.3) or (2.4) will turn up time and again in this course, so it will be convenient to introduce the following terminology

Definition 2.30. If $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are vectors in in $\mathbb{R}^m$ and $\alpha_1, \ldots, \alpha_n$ are scalars, a sum of the form

$$\alpha_1 \mathbf{a}_1 + \cdots + \alpha_n \mathbf{a}_n$$

is called a **linear combination** of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ with **weights** $\alpha_1, \ldots, \alpha_n$.

Summarising the previous discussion, we now have the following characterisation of consistency:

Theorem 2.31 (Consistency Theorem for Linear Systems)**.** *A linear system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ can be written as a linear combination of the column vectors of A.*

2.5 Elementary matrices and the Invertible Matrix Theorem

Using the reformulation of linear systems discussed in the previous section we shall now have another look at the process of solving them. Instead of performing elementary row operations we shall now view this process in terms of matrix multiplication. This will shed some light on both matrices and linear systems and will be useful for formulating and proving the main result of this chapter, the Invertible Matrix Theorem, which will be presented towards the end of this section. Before doing so, however, we shall consider the effect of multiplying both sides of a linear system in matrix form by an invertible matrix.

Lemma 2.32. *Let A be an $m \times n$ matrix and let $\mathbf{b} \in \mathbb{R}^m$. Suppose that M is an invertible $m \times m$ matrix. The following two systems are equivalent:*

$$A\mathbf{x} = \mathbf{b} \tag{2.5}$$

$$MA\mathbf{x} = M\mathbf{b} \tag{2.6}$$

Proof. Note that if $\mathbf{x}$ satisfies (2.5), then it clearly satisfies (2.6). Conversely, suppose that $\mathbf{x}$ satisfies (2.6), that is,

$$MA\mathbf{x} = M\mathbf{b}\,.$$

Since M is invertible, we may multiply both sides of the above equation by M^{-1} from the left to obtain

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

so $IA\mathbf{x} = I\mathbf{b}$, and hence $A\mathbf{x} = \mathbf{b}$, that is, $\mathbf{x}$ satisfies (2.5). □

We now come back to the idea outlined at the beginning of this section. It turns out that we can 'algebraize' the process of applying an elementary row operation to a matrix A by left-multiplying A by a certain type of matrix, defined as follows:

Definition 2.33. An **elementary matrix** of **type I** (respectively, **type II**, **type III**) is a matrix obtained by applying an elementary row operation of type I (respectively, type II, type III) to an identity matrix.

Example 2.34.

$$\text{type I: } E_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{(take } I_3 \text{ and swap rows 1 and 2)}$$

$$\text{type II: } E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \quad \text{(take } I_3 \text{ and multiply row 3 by 4)}$$

$$\text{type III: } E_3 = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{(take } I_3 \text{ and add 2 times row 3 to row 1)}$$

Let us now consider the effect of left-multiplying an arbitrary 3×3 matrix A in turn by each of the three elementary matrices given in the previous example.

Example 2.35. Let $A = (a_{ij})_{3\times 3}$ and let E_l $(l = 1, 2, 3)$ be defined as in the previous example. Then

$$E_1A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

$$E_2A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 4a_{31} & 4a_{32} & 4a_{33} \end{pmatrix},$$

$$E_3A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} + 2a_{31} & a_{12} + 2a_{32} & a_{13} + 2a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

You should now pause and marvel at the following observation: interchanging rows 1 and 2 of A produces E_1A, multiplying row 3 of A by 4 produces E_2A, and adding 2 times row 3 to row 1 of A produces E_3A.

This example should convince you of the truth of the following theorem, the proof of which will be omitted as it is straightforward, slightly lengthy and not particularly instructive.

Theorem 2.36. *If E is an $m \times m$ elementary matrix obtained from I by an elementary row operation, then left-multiplying an $m \times n$ matrix A by E has the effect of performing that same row operation on A.*

Slightly deeper is the following:

Theorem 2.37. *If E is an elementary matrix, then E is invertible and E^{-1} is an elementary matrix of the same type.*

Proof. The assertion follows from the previous theorem and the observation that an elementary row operation can be reversed by an elementary row operation of the same type. More precisely,

- if two rows of a matrix are interchanged, then interchanging them again restores the original matrix;
- if a row is multiplied by $\alpha \neq 0$, then multiplying the same row by $1/\alpha$ restores the original matrix;
- if α times row q has been added to row r, then adding $-\alpha$ times row q to row r restores the original matrix.

Now, suppose that E was obtained from I by a certain row operation. Then, as we just observed, there is another row operation of the same type that changes E back to I. Thus there is an elementary matrix F of the same type as E such that $FE = I$. A moment's thought shows that $EF = I$ as well, since E and F correspond to reverse operations. All in all, we have now shown that E is invertible and its inverse $E^{-1} = F$ is an elementary matrix of the same type. □

Example 2.38. Determine the inverses of the elementary matrices E_1, E_2, and E_3 in Example 2.34.

Solution. In order to transform E_1 into I we need to swap rows 1 and 2 of E_1. The elementary matrix that performs this feat is

$$E_1^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Similarly, in order to transform E_2 into I we need to multiply row 3 of E_2 by $\frac{1}{4}$. Thus

$$E_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix} .$$

Finally, in order to transform E_3 into I we need to add -2 times row 3 to row 1, and so

$$E_3^{-1} = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

□

Before we come to the main result of this chapter we need some more terminology:

Definition 2.39. A matrix B is **row equivalent** to a matrix A if there exists a finite sequence $E_1, E_2, \ldots, E_k$ of elementary matrices such that

$$B = E_k E_{k-1} \cdots E_1 A .$$

In other words, B is row equivalent to A if and only if B can be obtained from A by a finite number of row operations. In particular, two augmented matrices $(A|\mathbf{b})$ and $(B|\mathbf{c})$ are row equivalent if and only if $A\mathbf{x} = \mathbf{b}$ and $B\mathbf{x} = \mathbf{c}$ are equivalent systems.

The following properties of row equivalent matrices are easily established:

Fact 2.40. [1]

(a) A is row equivalent to itself;

(b) if A is row equivalent to B, then B is row equivalent to A;

(c) if A is row equivalent to B, and B is row equivalent to C, then A is row equivalent to C.

Property (b) follows from Theorem 2.37. Details of the proof of (a), (b), and (c) are left as an exercise.

We are now able to formulate and prove the first highlight of this module, a truly delightful characterisation of invertibility of matrices. More precisely, the following theorem provides three equivalent conditions for a matrix to be invertible. Later on in this course, we will encounter further equivalent conditions.

Before stating the theorem we recall that the **zero vector**, denoted by $\mathbf{0}$, is the column vector all of whose entries are zero.

Theorem 2.41 (Invertible Matrix Theorem)**.** *Let A be a square $n \times n$ matrix. The following are equivalent:*

(a) A is invertible;

(b) $A\mathbf{x} = \mathbf{0}$ has only the trivial solution;

(c) A is row equivalent to I;

(d) A is a product of elementary matrices.

Proof. We shall prove this theorem using a cyclic argument: we shall first show that (a) implies (b), then (b) implies (c), then (c) implies (d), and finally that (d) implies (a). This is a frequently used trick to show the logical equivalence of a list of assertions.

(a) $\Rightarrow$ (b): Suppose that A is invertible. If $\mathbf{x}$ satisfies $A\mathbf{x} = \mathbf{0}$, then

$$\mathbf{x} = I\mathbf{x} = (A^{-1}A)\mathbf{x} = A^{-1}\mathbf{0} = \mathbf{0}\,,$$

so the only solution of $A\mathbf{x} = \mathbf{0}$ is the trivial solution.

(b) $\Rightarrow$ (c): Use elementary row operations to bring the system $A\mathbf{x} = \mathbf{0}$ to the form $U\mathbf{x} = \mathbf{0}$, where U is in row echelon form. Since, by hypothesis, the solution of $A\mathbf{x} = \mathbf{0}$ and hence the solution of $U\mathbf{x} = \mathbf{0}$ is unique, there must be exactly n leading variables. Thus U is upper triangular with 1's on the diagonal, and hence, the reduced row echelon form of U is I. Thus A is row equivalent to I.

(c) $\Rightarrow$ (d): If A is row equivalent to I, then there is a sequence $E_1, \ldots, E_k$ of elementary matrices such that

$$A = E_k E_{k-1} \cdots E_1 I = E_k E_{k-1} \cdots E_1\,,$$

that is, A is a product of elementary matrices.

(d) $\Rightarrow$ (a). If A is a product of elementary matrices, then A must be invertible, since elementary matrices are invertible by Theorem 2.37 and since the product of invertible matrices is invertible by Theorem 2.18. $\square$

[1] In the language of MTH4104 (Introduction to Algebra) which some of you will have taken, these statements mean that 'row equivalence' is an equivalence relation.

An immediate consequence of the previous theorem is the following perhaps surprising result:

Corollary 2.42. *Let A and C be square matrices such that $CA = I$, then also $AC = I$; in particular both A and C are invertible with $C = A^{-1}$ and $A = C^{-1}$.*

Proof. By Exercise 1 on Coursework 3 it follows that A is invertible. But then C is invertible since

$$C = CI = CAA^{-1} = IA^{-1} = A^{-1}\,.$$

Furthermore, $C^{-1} = (A^{-1})^{-1} = A$ and $AC = C^{-1}C = I$. □

What is surprising about this result is the following: suppose we are given a square matrix A. If we want to check that A is invertible, then, by the definition of invertibility, we need to produce a matrix B such that $AB = I$ *and* $BA = I$. The above corollary tells us that if we have a candidate C for an inverse of A it is enough to check that *either* $AC = I$ *or* $CA = I$ in order to guarantee that A is invertible with inverse C. This is a non-trivial fact about matrices, which is often useful.

2.6 Gauss-Jordan inversion

The Invertible Matrix Theorem provides a simple method for inverting matrices. Recall that the theorem states (amongst other things) that if A is invertible, then A is row equivalent to I. Thus there is a sequence $E_1, \dots\ E_k$ of elementary matrices such that

$$E_k E_{k-1} \cdots E_1 A = I\,.$$

Multiplying both sides of the above equation by A^{-1} from the right yields

$$E_k E_{k-1} \cdots E_1 = A^{-1}\,,$$

that is,

$$E_k E_{k-1} \cdots E_1 I = A^{-1}\,.$$

Thus, the same sequence of elementary row operations that brings an invertible matrix to I, will bring I to A^{-1}. This gives a practical algorithm for inverting matrices, known as Gauss-Jordan inversion.

Note that in the following we use a slight generalisation of the augmented matrix notation. Given an $m \times n$ matrix A and an m-vector $\mathbf{b}$ we currently use $(A|\mathbf{b})$ to denote the $m \times (n+1)$ matrix consisting of A with $\mathbf{b}$ attached as an extra column to the right of A, and a vertical line in between them. Suppose now that B is an $m \times r$ matrix then we write $(A|B)$ for the $m \times (n + r)$ matrix consisting of A with B attached to the right of A, and a vertical line separating them.

Gauss-Jordan inversion

Bring the augmented matrix $(A|I)$ to reduced row echelon form. If A is row equivalent to I, then $(A|I)$ is row equivalent to $(I|A^{-1})$. Otherwise, A does not have an inverse.

Example 2.43. Show that

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 5 & 3 \\ 0 & 3 & 8 \end{pmatrix}$$

is invertible and compute A^{-1}.

Solution. Using Gauss-Jordan inversion we find

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 0 & 3 & 8 & 0 & 0 & 1 \end{array}\right) \sim_{R_2-2R_1} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 3 & -2 & 1 & 0 \\ 0 & 3 & 8 & 0 & 0 & 1 \end{array}\right)$$

$$\sim_{R_3-3R_2} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 3 & -2 & 1 & 0 \\ 0 & 0 & -1 & 6 & -3 & 1 \end{array}\right) \sim_{(-1)R_3} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 3 & -2 & 1 & 0 \\ 0 & 0 & 1 & -6 & 3 & -1 \end{array}\right)$$

$$\sim_{R_2-3R_3} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 16 & -8 & 3 \\ 0 & 0 & 1 & -6 & 3 & -1 \end{array}\right) \sim_{R_1-2R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -31 & 16 & -6 \\ 0 & 1 & 0 & 16 & -8 & 3 \\ 0 & 0 & 1 & -6 & 3 & -1 \end{array}\right).$$

Thus A is invertible (because it is row equivalent to I_3) and

$$A^{-1} = \begin{pmatrix} -31 & 16 & -6 \\ 16 & -8 & 3 \\ -6 & 3 & -1 \end{pmatrix}.$$

□

Chapter 3

Determinants

3.1 General definition of determinants

Let $A = (a_{ij})$ be a 2×2 matrix. Recall that the determinant of A was defined by

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}\,. \tag{3.1}$$

In other words, with every 2×2 matrix A it is possible to associate a scalar, called the determinant of A, which is given by a certain sum of products of the entries of A. The following fact was proved in Geometry I by a brute force calculation:

Fact 3.1. If A and B are 2×2 matrices then

(a) $\det(A) \neq 0$ if and only if A is invertible;

(b) $\det(AB) = \det(A)\det(B)$.

This fact reveals one of the main motivations to introduce this somewhat non-intuitive object: the determinant of a matrix allows us to decide whether a matrix is invertible or not.

In this chapter we introduce determinants for arbitrary square matrices, study some of their properties, and then prove the generalisation of the above fact for arbitrary square matrices.

Before giving the general definition of the determinant of an $n \times n$ matrix, let us recall the definition of 3×3 determinants given in Geometry I:

If $A = (a_{ij})$ is a 3×3 matrix, then its determinant is defined by

$$\begin{aligned} \det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}\,. \end{aligned} \tag{3.2}$$

Notice that the determinant of a 3×3 matrix A is given in terms of the determinants of certain 2×2 submatrices of A. In general, we shall see that the determinant of a 4×4 matrix is given in terms of the determinants of 3×3 submatrices, and so forth. Before stating the general definition we introduce a convenient short-hand:

Notation 3.2. For any square matrix A, let A_{ij} denote the submatrix formed by deleting the i-th row and the j-th column of A.

Warning: This notation differs from the one used in the course text

Example 3.3. If

$$A = \begin{pmatrix} 3 & 2 & 5 & -1 \\ -2 & 9 & 0 & 6 \\ 7 & -2 & -3 & 1 \\ 4 & -5 & 8 & -4 \end{pmatrix},$$

then

$$A_{23} = \begin{pmatrix} 3 & 2 & -1 \\ 7 & -2 & 1 \\ 4 & -5 & -4 \end{pmatrix}.$$

If we now define the determinant of a 1×1 matrix $A = (a_{ij})$ by $\det(A) = a_{11}$, we can recast (3.1) and (3.2) as follows:

- if $A = (a_{ij})_{2\times 2}$ then

$$\det(A) = a_{11}\det(A_{11}) - a_{21}\det(A_{21});$$

- if $A = (a_{ij})_{3\times 3}$ then

$$\det(A) = a_{11}\det(A_{11}) - a_{21}\det(A_{21}) + a_{31}\det(A_{31}).$$

This observation motivates the following recursive definition:

Definition 3.4. Let $A = (a_{ij})$ be an $n \times n$ matrix. The **determinant** of A, written $\det(A)$, is defined as follows:

- If $n = 1$, then $\det(A) = a_{11}$.
- If $n > 1$ then $\det(A)$ is the sum of n terms of the form $\pm a_{i1}\det(A_{i1})$, with plus and minus signs alternating, and where the entries $a_{11}, a_{21}, \ldots, a_{n1}$ are from the first column of A. In symbols:

$$\begin{aligned} \det(A) &= a_{11}\det(A_{11}) - a_{21}\det(A_{21}) + \cdots + (-1)^{n+1}a_{n1}\det(A_{n1}) \\ &= \sum_{i=1}^{n}(-1)^{i+1}a_{i1}\det(A_{i1}). \end{aligned}$$

Example 3.5. Compute the determinant of

$$A = \begin{pmatrix} 0 & 0 & 7 & -5 \\ -2 & 9 & 6 & -8 \\ 0 & 0 & -3 & 2 \\ 0 & 3 & -1 & 4 \end{pmatrix}.$$

Solution.

$$\begin{vmatrix} 0 & 0 & 7 & -5 \\ -2 & 9 & 6 & -8 \\ 0 & 0 & -3 & 2 \\ 0 & 3 & -1 & 4 \end{vmatrix} = -(-2)\begin{vmatrix} 0 & 7 & -5 \\ 0 & -3 & 2 \\ 3 & -1 & 4 \end{vmatrix} = 2\cdot 3\begin{vmatrix} 7 & -5 \\ -3 & 2 \end{vmatrix} = 2\cdot 3\cdot[7\cdot 2 - (-3)\cdot(-5)] = -6.$$

□

To state the next theorem, it will be convenient to write the definition of $\det(A)$ in a slightly different form.

Definition 3.6. Given a square matrix $A = (a_{ij})$, the (i,j)-**cofactor** of A is the number C_{ij} defined by

$$C_{ij} = (-1)^{i+j} \det(A_{ij}) .$$

Thus, the definition of $\det(A)$ reads

$$\det(A) = a_{11}C_{11} + a_{21}C_{21} + \cdots + a_{n1}C_{n1}.$$

This is called the **cofactor expansion down the first column of** A. There is nothing special about the first column, as the next theorem shows:

Theorem 3.7 (Cofactor Expansion Theorem)**.** *The determinant of an $n \times n$ matrix A can be computed by a cofactor expansion across any column or row. The expansion down the j-th column is*

$$\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

and the cofactor expansion across the i-th row is

$$\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} .$$

Although this theorem is fundamental for the development of determinants, we shall not prove it here, as it would lead to a rather lengthy digression.[1]

Before moving on, notice that the plus or minus sign in the (i,j)-cofactor depends on the position of a_{ij} in the matrix, regardless of a_{ij} itself. The factor $(-1)^{i+j}$ determines the following checkerboard pattern of signs

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

Example 3.8. Use a cofactor expansion across the second row to compute $\det(A)$, where

$$A = \begin{pmatrix} 4 & -1 & 3 \\ 0 & 0 & 2 \\ 1 & 0 & 7 \end{pmatrix} .$$

Solution.

$$\begin{aligned} \det(A) &= a_{21}C_{21} + a_{22}C_{22} + a_{23}C_{23} \\ &= (-1)^{2+1}a_{21}\det(A_{21}) + (-1)^{2+2}a_{22}\det(A_{22}) + (-1)^{2+3}a_{23}\det(A_{23}) \\ &= -0\begin{vmatrix} -1 & 3 \\ 0 & 7 \end{vmatrix} + 0\begin{vmatrix} 4 & 3 \\ 1 & 7 \end{vmatrix} - 2\begin{vmatrix} 4 & -1 \\ 1 & 0 \end{vmatrix} \\ &= -2[4 \cdot 0 - 1 \cdot (-1)] = -2 . \end{aligned}$$

□

[1] If not knowing the proof causes you sleepless nights, ask me and I will point you towards one.

Example 3.9. Compute $\det(A)$, where

$$A = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ -2 & 5 & 0 & 0 & 0 \\ 9 & -6 & 4 & -1 & 3 \\ 2 & 4 & 0 & 0 & 2 \\ 8 & 3 & 1 & 0 & 7 \end{pmatrix}.$$

Solution. Notice that all entries but the first of row 1 are 0. Thus it will shorten our labours if we expand across the first row:

$$\det(A) = 3 \begin{vmatrix} 5 & 0 & 0 & 0 \\ -6 & 4 & -1 & 3 \\ 4 & 0 & 0 & 2 \\ 3 & 1 & 0 & 7 \end{vmatrix}.$$

Again it is advantageous to expand this 4×4 determinant across the first row:

$$\det(A) = 3 \cdot 5 \cdot \begin{vmatrix} 4 & -1 & 3 \\ 0 & 0 & 2 \\ 1 & 0 & 7 \end{vmatrix}.$$

We have already computed the value of the above 3×3 determinant in the previous example and found it to be equal to -2. Thus $\det(A) = 3 \cdot 5 \cdot (-2) = -30$. □

Notice that the matrix in the previous example was almost lower triangular. The method of this example is easily generalised to prove the following theorem:

Theorem 3.10. *If A is either an upper or a lower triangular matrix, then $\det(A)$ is the product of the diagonal entries of A.*

3.2 Properties of determinants

We saw time and again in this module that elementary row operations play a fundamental role in matrix theory. It is only natural to enquire how $\det(A)$ behaves when an elementary row operation is applied to A.

Theorem 3.11. *Let A be a square matrix.*

(a) If two rows of A are interchanged to produce B, then $\det(B) = -\det(A)$.

(b) If one row of A is multiplied by α to produce B, then $\det(B) = \alpha \det(A)$.

(c) If a multiple of one row of A is added to another row to produce a matrix B then $\det(B) = \det(A)$.

Proof. These assertions follow from a slightly stronger result to be proved later in this chapter (see Theorem 3.21). □

Example 3.12.

(a) $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = - \begin{vmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{vmatrix}$ by (a) of the previous theorem.

(b) $\begin{vmatrix} 0 & 1 & 2 \\ 3 & 12 & 9 \\ 1 & 2 & 1 \end{vmatrix} = 3 \begin{vmatrix} 0 & 1 & 2 \\ 1 & 4 & 3 \\ 1 & 2 & 1 \end{vmatrix}$ by (b) of the previous theorem.

(c) $\begin{vmatrix} 3 & 1 & 0 \\ 4 & 2 & 9 \\ 0 & -2 & 1 \end{vmatrix} = \begin{vmatrix} 3 & 1 & 0 \\ 7 & 3 & 9 \\ 0 & -2 & 1 \end{vmatrix}$ by (c) of the previous theorem.

The following examples show how to use the previous theorem for the effective computation of determinants:

Example 3.13. Compute

$$\begin{vmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ -6 & 7 & -7 & 4 \\ -5 & -8 & 0 & 9 \end{vmatrix}.$$

Solution. Perhaps the easiest way to compute this determinant is to spot that when adding two times row 1 to row 3 we get two identical rows, which, by another application of the previous theorem, implies that the determinant is zero:

$$\begin{vmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ -6 & 7 & -7 & 4 \\ -5 & -8 & 0 & 9 \end{vmatrix} \overset{R_3 + 2R_1}{=} \begin{vmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 5 & -3 & -6 \\ -5 & -8 & 0 & 9 \end{vmatrix}$$

$$\overset{R_3 - R_2}{=} \begin{vmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 0 & 0 & 0 \\ -5 & -8 & 0 & 9 \end{vmatrix} = 0,$$

by a cofactor expansion across the third row. □

Example 3.14. Compute $\det(A)$, where

$$A = \begin{pmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ -2 & -5 & 4 & -2 \end{pmatrix}.$$

Solution. Here we see that the first column already has two zero entries. Using the previous theorem we can introduce another zero in this column by adding row 2 to row 4. Thus

$$\det(A) = \begin{vmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ -2 & -5 & 4 & -2 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ 0 & 0 & -3 & 1 \end{vmatrix}.$$

If we now expand down the first column we see that

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

The 3×3 determinant above can be further simplified by subtracting 3 times row 1 from row 2. Thus

$$\det(A) = -2 \begin{vmatrix} 1 & 2 & -1 \\ 0 & 0 & 5 \\ 0 & -3 & 1 \end{vmatrix}.$$

Finally we notice that the above determinant can be brought to triangular form by swapping row 2 and row 3, which changes the sign of the determinant by the previous theorem. Thus

$$\det(A) = (-2) \cdot (-1) \begin{vmatrix} 1 & 2 & -1 \\ 0 & -3 & 1 \\ 0 & 0 & 5 \end{vmatrix} = (-2) \cdot (-1) \cdot 1 \cdot (-3) \cdot 5 = -30\,,$$

by Theorem 3.10. □

We are now able to prove the first important result about determinants. It allows us to decide whether a matrix is invertible or not by computing its determinant. It will play an important role in later chapters.

Theorem 3.15. *A matrix A is invertible if and only if $\det(A) \neq 0$.*

Proof. Bring A to row echelon form U (which is then necessarily upper triangular). Since we can achieve this using elementary row operations, and since, in the process we only ever multiply a row by a non-zero scalar $\det(A) = \gamma \det(U)$ for some γ with $\gamma \neq 0$, by Theorem 3.11. If A is invertible, then $\det(U) = 1$, since U is upper triangular with 1's on the diagonal, and hence $\det(A) = \gamma \det(U) \neq 0$. Otherwise, at least one diagonal entry of U is zero, so $\det(U) = 0$, and hence $\det(A) = \gamma \det(U) = 0$. □

Definition 3.16. A square matrix A is called **singular** if $\det(A) = 0$. Otherwise it is said to be **nonsingular**.

Corollary 3.17. *A matrix is invertible if and only if it is nonsingular*

Our next result shows what effect transposing a matrix has on its determinant:

Theorem 3.18. *If A is an $n \times n$ matrix, then $\det(A) = \det(A^T)$.*

Proof. The proof is by induction on n (that is, the size of A).[2] The theorem is obvious for $n = 1$. Suppose now that it has already been proved for $k \times k$ matrices for some integer k. Our aim now is to show that the assertion of the theorem is true for $(k+1) \times (k+1)$ matrices as well. Let A be a $(k+1) \times (k+1)$ matrix. Note that the (i,j)-cofactor of A equals the (i,j)-cofactor of A^T, because the cofactors involve $k \times k$ determinants only, for which we assumed that the assertion of the theorem holds. Hence

$$\begin{gathered} \text{cofactor expansion of } \det(A) \text{ across first row} \\ = \text{cofactor expansion of } \det(A^T) \text{ down first column} \end{gathered}$$

so $\det(A) = \det(A^T)$.

Let's summarise: the theorem is true for 1×1 matrices, and the truth of the theorem for $k \times k$ matrices for some k implies the truth of the theorem for $(k+1) \times (k+1)$ matrices.

[2] If you have never encountered this method of proof, don't despair! Simply read through the following argument. The last paragraph explains the underlying idea of this method.

Thus, the theorem must be true for 2×2 matrices (choose $k = 1$); but since we now know that it is true for 2×2 matrices, it must be true for 3×3 matrices as well (choose $k = 2$); continuing with this process, we see that the theorem must be true for matrices of arbitrary size. □

By the previous theorem, each statement of the theorem on the behaviour of determinants under row operations (Theorem 3.11) is also true if the word 'row' is replaced by 'column', since a row operation on A^T amounts to a column operation on A.

Theorem 3.19. *Let A be a square matrix.*

(a) If two columns of A are interchanged to produce B, then $\det(B) = -\det(A)$.

(b) If one column of A is multiplied by α to produce B, then $\det(B) = \alpha \det(A)$.

(c) If a multiple of one column of A is added to another column to produce a matrix B then $\det(B) = \det(A)$.

Example 3.20. Find $\det(A)$ where

$$A = \begin{pmatrix} 1 & 3 & 4 & 8 \\ -1 & 2 & 1 & 9 \\ 2 & 5 & 7 & 0 \\ 3 & -4 & -1 & 5 \end{pmatrix}.$$

Solution. Adding column 1 to column 2 gives

$$\det(A) = \begin{vmatrix} 1 & 3 & 4 & 8 \\ -1 & 2 & 1 & 9 \\ 2 & 5 & 7 & 0 \\ 3 & -4 & -1 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 4 & 4 & 8 \\ -1 & 1 & 1 & 9 \\ 2 & 7 & 7 & 0 \\ 3 & -1 & -1 & 5 \end{vmatrix}.$$

Now subtracting column 3 from column 2 the determinant is seen to vanish by a cofactor expansion down column 2.

$$\det(A) = \begin{vmatrix} 1 & 0 & 4 & 8 \\ -1 & 0 & 1 & 9 \\ 2 & 0 & 7 & 0 \\ 3 & 0 & -1 & 5 \end{vmatrix} = 0\,.$$

□

Our next aim is to prove that determinants are multiplicative, that is, $\det(AB) = \det(A)\det(B)$ for any two square matrices A and B of the same size. We start by establishing a baby-version of this result, which, at the same time, proves the theorem on the behaviour of determinants under row operations stated earlier (see Theorem 3.11).

Theorem 3.21. *If A is an $n \times n$ matrix and E an elementary $n \times n$ matrix, then*

$$\det(EA) = \det(E)\det(A)$$

with

$$\det(E) = \begin{cases} -1 & \text{if } E \text{ is of type I (interchanging two rows)} \\ \alpha & \text{if } E \text{ is of type II (multiplying a row by } \alpha) \\ 1 & \text{if } E \text{ is of type III (adding a multiple of one row to another)} \end{cases}.$$

Proof. By induction on the size of A. The case where A is a 2×2 matrix is an easy exercise (see Exercise 1, Coursework 4). Suppose now that the theorem has been verified for determinants of $k\times k$ matrices for some k with $k\geq 2$. Let A be $(k+1)\times(k+1)$ matrix and write $B=EA$. Expand $\det(EA)$ across a row that is unaffected by the action of E on A, say, row i. Note that B_{ij} is obtained from A_{ij} by the same type of elementary row operation that E performs on A. But since these matrices are only $k\times k$, our hypothesis implies that

$$\det(B_{ij}) = r\det(A_{ij})\,,$$

where $r=-1,\alpha,1$ depending on the nature of E.

Now by a cofactor expansion across row i

$$\begin{aligned}\det(EA)=\det(B) &= \sum_{j=1}^{k+1} a_{ij}(-1)^{i+j}\det(B_{ij})\\ &= \sum_{j=1}^{k+1} a_{ij}(-1)^{i+j} r\det(A_{ij})\\ &= r\det(A)\,.\end{aligned}$$

In particular, taking $A=I_{k+1}$ we see that $\det(E)=-1,\alpha,1$ depending on the nature of E.

To summarise: the theorem is true for 2×2 matrices and the truth of the theorem for $k\times k$ matrices for some $k\geq 2$ implies the truth of the theorem for $(k+1)\times(k+1)$ matrices. By the principle of induction the theorem is true for matrices of any size. □

Using the previous theorem we are now able to prove the second important result of this chapter:

Theorem 3.22. *If A and B are square matrices of the same size, then*

$$\det(AB)=\det(A)\det(B)\,.$$

Proof. Case I: If A is not invertible, then neither is AB (for otherwise $A(B(AB)^{-1})=I$, which by the corollary to the Invertible Matrix Theorem (Corollary 2.42), would force A to be invertible). Thus, by Theorem 3.15,

$$\det(AB)=0=0\cdot\det(B)=\det(A)\det(B)\,.$$

Case II: If A is invertible, then by the Invertible Matrix Theorem A is a product of elementary matrices, that is, there exist elementary matrices $E_1,\ldots,E_k$, such that

$$A=E_kE_{k-1}\cdots E_1\,.$$

For brevity, write $|A|$ for $\det(A)$. Then, by the previous theorem,

$$\begin{aligned}|AB| &= |E_k\cdots E_1B| = |E_k||E_{k-1}\cdots E_1B| = \ldots\\ &= |E_k|\cdots|E_1||B| = \ldots = |E_k\cdots E_1||B|\\ &= |A||B|\,.\end{aligned}$$

□

3.3 Cramer's Rule and a formula for A^{-1}

In the following, we use a shorthand to identify a matrix by its columns. Let A be an $m \times n$ matrix. If $\mathbf{a}_j \in \mathbb{R}^m$ is the j-th column vector of A, then we write

$$A = (\mathbf{a}_1 \dots \mathbf{a}_n) .$$

Note that if B is an $l \times m$ matrix then, by the definition of matrix multiplication,

$$BA = (B\mathbf{a}_1 \dots B\mathbf{a}_n) ,$$

that is, the j-th column of BA is $B\mathbf{a}_j$.

Cramer's Rule is a curious formula that allows us to write down the solution for certain $n \times n$ systems in terms of quotients of two determinants. Before stating it we need some more notation.

For any $n \times n$ matrix A and any $\mathbf{b} \in \mathbb{R}^n$ write $A_i(\mathbf{b})$ for the matrix obtained from A by replacing column i by $\mathbf{b}$, that is,

$$A_i(\mathbf{b}) = (\mathbf{a}_1 \dots \underset{\text{col } i}{\mathbf{b}} \dots \mathbf{a}_n) .$$

Theorem 3.23 (Cramer's Rule)**.** *Let A be an invertible $n \times n$ matrix. For any $\mathbf{b} \in \mathbb{R}^n$, the unique solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}$ has entries given by*

$$x_i = \frac{\det(A_i(\mathbf{b}))}{\det(A)} \quad \textit{for } i = 1, \dots, n .$$

Proof. Let $\mathbf{a}_1 \dots, \mathbf{a}_n$ be the columns of A, and $\mathbf{e}_1, \dots, \mathbf{e}_n$ the columns of the $n \times n$ identity matrix I. Then

$$\begin{aligned} AI_i(\mathbf{x}) &= A(\mathbf{e}_1 \dots \mathbf{x} \dots \mathbf{e}_n) \\ &= (A\mathbf{e}_1 \dots A\mathbf{x} \dots A\mathbf{e}_n) \\ &= (\mathbf{a}_1 \dots \mathbf{b} \dots \mathbf{a}_n) \\ &= A_i(\mathbf{b}) , \end{aligned}$$

and hence

$$\det(A) \det(I_i(\mathbf{x})) = \det(A_i(\mathbf{b})) .$$

But $\det(I_i(\mathbf{x})) = x_i$ by a cofactor expansion across row i, so

$$x_i = \frac{\det(A_i(\mathbf{b}))}{\det(A)} ,$$

since $\det(A) \neq 0$. □

Example 3.24. Use Cramer's Rule to solve the system

$$\begin{array}{rcrl} 3x_1 & - & 2x_2 & = 6 \\ -5x_1 & + & 4x_2 & = 8 \end{array} .$$

Solution. Write the system as $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{pmatrix} 3 & -2 \\ -5 & 4 \end{pmatrix} , \quad \mathbf{b} = \begin{pmatrix} 6 \\ 8 \end{pmatrix} .$$

Then

$$A_1(\mathbf{b}) = \begin{pmatrix} 6 & -2 \\ 8 & 4 \end{pmatrix}, \quad A_2(\mathbf{b}) = \begin{pmatrix} 3 & 6 \\ -5 & 8 \end{pmatrix},$$

and Cramer's Rule gives

$$x_1 = \frac{\det(A_1(\mathbf{b}))}{\det(A)} = \frac{40}{2} = 20\,,$$
$$x_2 = \frac{\det(A_2(\mathbf{b}))}{\det(A)} = \frac{54}{2} = 27\,.$$

□

Cramer's Rule is not really useful for practical purposes (except for very small systems), since evaluation of determinants is time consuming when the system is large. For 3×3 systems and larger, you are better off using Gaussian elimination. Apart from its intrinsic beauty, its main strength is as a theoretical tool. For example, it allows you to study how sensitive the solution of $A\mathbf{x} = \mathbf{b}$ is to a change in an entry in A or $\mathbf{b}$.

As an application of Cramer's Rule, we shall now derive an explicit formula for the inverse of a matrix. Before doing so we shall have another look at the process of inverting a matrix. Again, denote the columns of I_n by $\mathbf{e}_1, \ldots, \mathbf{e}_n$. The Gauss-Jordan inversion process bringing $(A|I)$ to $(I|A^{-1})$ can be viewed as solving the n systems

$$A\mathbf{x} = \mathbf{e}_1\,, \quad A\mathbf{x} = \mathbf{e}_2\,, \quad \ldots \quad A\mathbf{x} = \mathbf{e}_n\,.$$

Thus the j-th column of A^{-1} is the solution of

$$A\mathbf{x} = \mathbf{e}_j\,,$$

and the i-th entry of $\mathbf{x}$ is the (i,j)-entry of A^{-1}. By Cramer's rule

$$(i,j)\text{-entry of } A^{-1} = x_i = \frac{\det(A_i(\mathbf{e}_j))}{\det(A)}\,. \tag{3.3}$$

A cofactor expansion down column i of $A_i(\mathbf{e}_j))$ shows that

$$\det(A_i(\mathbf{e}_j)) = (-1)^{i+j}\det(A_{ji}) = C_{ji}\,,$$

where C_{ji} is the (j,i)-cofactor of A. Thus, by (3.3), the (i,j)-entry of A^{-1} is the cofactor C_{ji} divided by $\det(A)$ (note that the order of the indices is reversed!). Thus

$$A^{-1} = \frac{1}{\det(A)} \underbrace{\begin{pmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_n \end{pmatrix}}_{=\operatorname{adj}(A)}\,. \tag{3.4}$$

The matrix of cofactors on the right of (3.4) is called the **adjugate** of A, and is denoted by $\operatorname{adj}(A)$. The following theorem is simply a restatement of (3.4):

Theorem 3.25 (Inverse Formula)**.** *Let A be an invertible matrix. Then*

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

Example 3.26. Find the inverse of the following matrix using the Inverse Formula

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

Proof. First we need to calculate the 9 cofactors of A:

$$C_{11} = + \begin{vmatrix} -6 & 0 \\ 4 & -3 \end{vmatrix} = 18 , \quad C_{12} = - \begin{vmatrix} -2 & 0 \\ 1 & -3 \end{vmatrix} = -6 , \quad C_{13} = + \begin{vmatrix} -2 & -6 \\ 1 & 4 \end{vmatrix} = -2 ,$$
$$C_{21} = - \begin{vmatrix} 3 & -1 \\ 4 & -3 \end{vmatrix} = 5 , \quad C_{22} = + \begin{vmatrix} 1 & -1 \\ 1 & -3 \end{vmatrix} = -2 , \quad C_{23} = - \begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} = -1 ,$$
$$C_{31} = + \begin{vmatrix} 3 & -1 \\ -6 & 0 \end{vmatrix} = -6 , \quad C_{32} = - \begin{vmatrix} 1 & -1 \\ -2 & 0 \end{vmatrix} = 2 , \quad C_{33} = + \begin{vmatrix} 1 & 3 \\ -2 & -6 \end{vmatrix} = 0 .$$

Thus

$$\operatorname{adj}(A) = \begin{pmatrix} 18 & 5 & -6 \\ -6 & -2 & 2 \\ -2 & -1 & 0 \end{pmatrix} ,$$

and since $\det(A) = 2$, we have

$$A^{-1} = \begin{pmatrix} 9 & \frac{5}{2} & -3 \\ -3 & -1 & 1 \\ -1 & -\frac{1}{2} & 0 \end{pmatrix} .$$

□

Note that the above calculations are just as laborious as if we had used the Gauss-Jordan inversion process to compute A^{-1}. As with Cramer's Rule, the deceptively neat formula for the inverse is not useful if you want to invert larger matrices. As a rule, for matrices larger than 3×3 the Gauss-Jordan inversion algorithm is much faster.

Chapter 4

Vector Spaces

4.1 Definition and examples

In this chapter, we will study abstract vector spaces. Roughly speaking a vector space is a mathematical structure on which an operation of addition and an operation of scalar multiplication is defined, and we require these operations to obey a number of algebraic rules. We have already encountered examples of vector spaces in this module. Recall that $\mathbb{R}^n$ is the collection of all n-vectors. On $\mathbb{R}^n$ two operations were defined:

- *addition*: if
$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n,$$
then $\mathbf{x} + \mathbf{y}$ is the n-vector given by
$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

- *scalar multiplication*: if
$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \quad \text{and} \quad \alpha \text{ is a scalar}$$
then $\alpha\mathbf{x}$ is the n-vector given by
$$\alpha\mathbf{x} = \begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix}.$$

After these operations were defined, it turned out that they satisfy a number of rules (see Theorem 2.7).[1] We are now going to turn this process on its head. That is, we start from a set on which two operations are defined, we *postulate* that these operations satisfy certain rules, and we call the resulting structure a 'vector space':

[1] Go back and look them up!

Definition 4.1. A **vector space** is a non-empty set V on which are defined two operations, called *addition* and *scalar multiplication*, such that the following axioms hold for all $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ in V and all scalars α, β:

(C1) the sum of $\mathbf{u}$ and $\mathbf{v}$, denoted by $\mathbf{u}+\mathbf{v}$, is in V;

(C2) the scalar multiple of $\mathbf{u}$ by α, denoted by $\alpha\mathbf{u}$, is in V;

(A1) $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$;

(A2) $\mathbf{u}+(\mathbf{v}+\mathbf{w})=(\mathbf{u}+\mathbf{v})+\mathbf{w}$;

(A3) there is an element $\mathbf{0}$ in V such that $\mathbf{u}+\mathbf{0}=\mathbf{u}$;

(A4) for each $\mathbf{u}$ in V there is an element $-\mathbf{u}$ in V such that $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$;

(A5) $\alpha(\mathbf{u}+\mathbf{v})=\alpha\mathbf{u}+\alpha\mathbf{v}$;

(A6) $(\alpha+\beta)\mathbf{u}=\alpha\mathbf{u}+\beta\mathbf{u}$;

(A7) $(\alpha\beta)\mathbf{u}=\alpha(\beta\mathbf{u})$;

(A8) $1\mathbf{u}=\mathbf{u}$.

We will refer to V as the universal set for the vector space. Its elements are called **vectors**, and we usually write them using bold letters $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, etc.

The term 'scalar' will usually refer to a real number, although later on we will sometimes allow scalars to be complex numbers. To distinguish these cases we will use the term **real vector space** (if the scalars are real numbers) or **complex vector space** (if the scalars are complex numbers). For the moment, however, we will only consider real vector spaces.

Note that in the above definition the axioms (C1) and (C2), known as closure axioms, simply state that the two operations produce values in V, that is, they are *bona fide* operations on V. The other eight axioms, also known as the classical vector space axioms, stipulate how the two operations interact.

Let's have a look at some examples:

Example 4.2. Let $\mathbb{R}^{m\times n}$ denote the set of all $m\times n$ matrices. Define addition and scalar multiplication of matrices in the usual way. Then $\mathbb{R}^{n\times m}$ is a vector space by Theorem 2.7.

Example 4.3. Let P_n denote the set of all polynomials with real coefficients of degree less than n. Thus, an element $\mathbf{p}$ in P_n is of the form

$$\mathbf{p}(t)=a_0+a_1t+a_2t^2+\cdots+a_nt^n\,,$$

where the coefficients $a_0,\ldots,a_n$ and the variable t are real numbers.

Define addition and scalar multiplication on P_n as follows: if $\mathbf{q}\in P_n$ is given by

$$\mathbf{q}(t)=b_0+b_1t+b_2t^2+\cdots+b_nt^n\,,$$

$\mathbf{p}$ is as above and α a scalar, then

- $\mathbf{p}+\mathbf{q}$ is the polynomial

$$(\mathbf{p}+\mathbf{q})(t)=(a_0+b_0)+(a_1+b_1)t+\cdots+(a_n+b_n)t^n$$

- $\alpha\mathbf{p}$ is the polynomial

$$(\alpha\mathbf{p})(t) = (\alpha a_0) + (\alpha a_1)t + \cdots + (\alpha a_n)t^n\,.$$

Note that (C1) and (C2) clearly hold, since if $\mathbf{p}, \mathbf{q} \in P_n$ and α is a scalar, then $\mathbf{p}+\mathbf{q}$ and $\alpha\mathbf{p}$ are again polynomials of degree less than n. Axiom (A1) holds since if $\mathbf{p}$ and $\mathbf{q}$ are as above, then

$$\begin{aligned}(\mathbf{p}+\mathbf{q})(t) &= (a_0+b_0) + (a_1+b_1)t + \cdots + (a_n+b_n)t^n \\ &= (b_0+a_0) + (b_1+a_1)t + \cdots + (b_n+a_n)t^n \\ &= (\mathbf{q}+\mathbf{p})(t)\end{aligned}$$

so $\mathbf{p}+\mathbf{q} = \mathbf{q}+\mathbf{p}$. A similar calculation shows that (A2) holds. Axiom (A3) holds if we let $\mathbf{0}$ be the zero polynomial, that is

$$\mathbf{0}(t) = 0 + 0\cdot t + \cdots + 0\cdot t^n\,,$$

since then $(\mathbf{p}+\mathbf{0})(t) = \mathbf{p}(t)$, that is, $\mathbf{p}+\mathbf{0} = \mathbf{p}$. Axiom (A4) holds if, given $\mathbf{p} \in P_n$ we set $-\mathbf{p} = (-1)\mathbf{p}$, since then

$$(\mathbf{p}+(-\mathbf{p}))(t) = (a_0-a_0) + (a_1-a_1)t + \cdots + (a_n-a_n)t^n = \mathbf{0}(t)\,,$$

that is $\mathbf{p}+(-\mathbf{p}) = \mathbf{0}$. The remaining axioms are easily verified as well, using familiar properties of real numbers.

Example 4.4. Let $C[a,b]$ denote the set of all real-valued functions that are defined and continuous on the closed interval $[a,b]$. For $\mathbf{f}, \mathbf{g} \in C[a,b]$ and α a scalar, define $\mathbf{f}+\mathbf{g}$ and $\alpha\mathbf{f}$ *pointwise*, that is, by

$$(\mathbf{f}+\mathbf{g})(t) = \mathbf{f}(t) + \mathbf{g}(t) \quad \text{for all } t \in [a,b]$$

$$(\alpha\mathbf{f})(t) = \alpha\mathbf{f}(t) \quad \text{for all } t \in [a,b]$$

Equipped with these operations, $C[a,b]$ is a vector space. The closure axiom (C1) holds because the sum of two continuous functions on $[a,b]$ is continuous on $[a,b]$, and (C2) holds because a constant times a continuous function on $[a,b]$ is again continuous on $[a,b]$. Axiom (A1) holds as well, since for all $t \in [a,b]$

$$(\mathbf{f}+\mathbf{g})(t) = \mathbf{f}(t) + \mathbf{g}(t) = \mathbf{g}(t) + \mathbf{f}(t) = (\mathbf{g}+\mathbf{f})(t)\,,$$

so $\mathbf{f}+\mathbf{g} = \mathbf{g}+\mathbf{f}$. Axiom (A3) is satisfied if we let $\mathbf{0}$ be the zero function,

$$\mathbf{0}(t) = 0 \quad \text{for all } t \in [a,b]\,,$$

since then

$$(\mathbf{f}+\mathbf{0})(t) = \mathbf{f}(t) + \mathbf{0}(t) = \mathbf{f}(t) + 0 = \mathbf{f}(t)\,,$$

so $\mathbf{f}+\mathbf{0} = \mathbf{f}$. Axiom (A4) holds if, given $\mathbf{f} \in C[a,b]$, we let $-\mathbf{f}$ be the function

$$(-\mathbf{f})(t) = -\mathbf{f}(t) \quad \text{for all } t \in [a,b],$$

since then

$$(\mathbf{f}+(-\mathbf{f}))(t) = \mathbf{f}(t) + (-\mathbf{f})(t) = \mathbf{f}(t) - \mathbf{f}(t) = 0 = \mathbf{0}(t)\,,$$

that is, $\mathbf{f}+(-\mathbf{f}) = \mathbf{0}$. We leave it as an exercise to verify the remaining axioms.

We shall now derive a number of elementary properties of vector spaces.

Theorem 4.5. *If V is a vector space and $\mathbf{u}$ and $\mathbf{v}$ are elements in V, then*

(a) $0\mathbf{u} = \mathbf{0}$;

(b) if $\mathbf{u} + \mathbf{v} = \mathbf{0}$ *then* $\mathbf{v} = -\mathbf{u}$;[2]

(c) $(-1)\mathbf{u} = -\mathbf{u}$.

Proof. (a) We start by observing that

$$\mathbf{u} \overset{(A8)}{=} 1\mathbf{u} = (0+1)\mathbf{u} \overset{(A6)}{=} 0\mathbf{u} + 1\mathbf{u} \overset{(A8)}{=} 0\mathbf{u} + \mathbf{u}\,. \tag{4.1}$$

Now, by (A4), there is an element $-\mathbf{u} \in V$ such that

$$\mathbf{u} + (-\mathbf{u}) = \mathbf{0}\,. \tag{4.2}$$

Thus

$$\mathbf{0} \overset{(4.2)}{=} \mathbf{u} + (-\mathbf{u}) \overset{(4.1)}{=} (0\mathbf{u} + \mathbf{u}) + (-\mathbf{u}) \overset{(A2)}{=} 0\mathbf{u} + (\mathbf{u} + (-\mathbf{u})) \overset{(4.2)}{=} 0\mathbf{u} + \mathbf{0} \overset{(A3)}{=} 0\mathbf{u}\,.$$

(b) Suppose that $\mathbf{u} + \mathbf{v} = \mathbf{0}$. Then

$$\begin{aligned} -\mathbf{u} \overset{(A3)}{=} -\mathbf{u} + \mathbf{0} = -\mathbf{u} + (\mathbf{u} + \mathbf{v}) &\overset{(A2)}{=} (-\mathbf{u} + \mathbf{u}) + \mathbf{v} \\ &\overset{(A1)}{=} (\mathbf{u} + (-\mathbf{u})) + \mathbf{v} \overset{(A4)}{=} \mathbf{0} + \mathbf{v} \overset{(A1)}{=} \mathbf{v} + \mathbf{0} \overset{(A3)}{=} \mathbf{v}\,. \end{aligned}$$

(c) Notice that

$$\mathbf{0} \overset{(a)}{=} 0\mathbf{u} = (1 + (-1))\mathbf{u} \overset{(A6)}{=} 1\mathbf{u} + (-1)\mathbf{u} \overset{(A8)}{=} \mathbf{u} + (-1)\mathbf{u}\,,$$

so, by (b), we conclude that $(-1)\mathbf{u} = -\mathbf{u}$. □

4.2 Subspaces

Given a vector space V, a 'subspace' of V is, roughly speaking, a subset of V that inherits the vector space structure from V, and can thus be considered as a vector space in its own right. One of the main motivations to consider such 'substructures' of vector spaces, is the following. As you might have noticed, it can be frightfully tedious to check whether a given set, call it H, is a vector space. Suppose that we know that H is a subset of a larger set V equipped with two operations (addition and scalar multiplication), for which we have already checked that the vector space axioms are satisfied. Now, in order for H to be the universal set of a vector space equipped with the operations of addition and scalar multiplication inherited from V, the set H should certainly be closed under addition and scalar multiplication (so that (C1) and (C2) are satisfied). Somewhat surprisingly, checking these two axioms is enough in order for H to be a vector space in its own right, as we shall see shortly. To summarise: if H is a subset of a vector space V, and if H is closed under addition and scalar multiplication, then H is a vector space in its own right. So instead of having to check 10 axioms, we only need to check two in this case. Let's cast these observations into the following definition:

[2] In the language of MTH4104 (Introduction to Algebra) this statement says that the additive inverse is unique.

Definition 4.6. Let H be a nonempty subset of a vector space V. Suppose that H satisfies the following two conditions:

(i) if $\mathbf{u}, \mathbf{v} \in H$, then $\mathbf{u} + \mathbf{v} \in H$;

(ii) if $\mathbf{u} \in H$ and α is a scalar, then $\alpha\mathbf{u} \in H$.

Then H is said to be a **subspace** of V.

Theorem 4.7. *Let H be a subspace of a vector space V. Then H with addition and scalar multiplication inherited from V is a vector space in its own right.*

Proof. Clearly, by definition of a subspace, (C1) and (C2) are satisfied. Axioms (A3) and (A4) follow from Theorem 4.5 and condition (ii) of the definition of a subspace. The remaining axioms are valid for any elements in V, so, in particular, they are valid for any elements in H as well. □

Remark 4.8. If V is a vector space, then $\{\mathbf{0}\}$ and V are clearly subspaces of V. All other subspaces are said to be **proper subspaces** of V. We call $\{\mathbf{0}\}$ the **zero subspace** of V.

Let's have a look at some more concrete examples:

Example 4.9. Show that the following are subspaces of $\mathbb{R}^3$:

(a) $L = \left\{ (r, s, t,)^T \,\middle|\, r, s, t \in \mathbb{R} \text{ and } r = s = t \right\}$;[3]

(b) $P = \left\{ (r, s, t,)^T \,\middle|\, r, s, t \in \mathbb{R} \text{ and } r - s + 3t = 0 \right\}$.

Solution. (a) Notice that an arbitrary element in L is of the form $r(1, 1, 1)^T$ for some real number r. Thus, in particular, L is not empty, since $(0, 0, 0)^T \in L$. In order to check that L is a subspace of $\mathbb{R}^3$ we need to check that conditions (i) and (ii) of Definition 4.6 are satisfied.

We start with condition (i). Let $\mathbf{x}_1$ and $\mathbf{x}_2$ belong to L. Then $\mathbf{x}_1 = r_1(1, 1, 1)^T$ and $\mathbf{x}_2 = r_2(1, 1, 1)^T$ for some real numbers r_1 and r_2, so

$$\mathbf{x}_1 + \mathbf{x}_2 = r_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + r_2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = (r_1 + r_2) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in L\,.$$

Thus condition (i) holds.

We now check condition (ii). Let $\mathbf{x} \in L$ and let α be a real number. Then $\mathbf{x} = r(1, 1, 1)^T$ for some real number $r \in \mathbb{R}$, so

$$\alpha\mathbf{x} = \alpha r \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in L\,.$$

Thus condition (ii) holds.

[3] In order to save paper, hence trees and thus do our bit to prevent climate change, we shall sometimes write n-vectors $\mathbf{x} \in \mathbb{R}^n$ in the form $(x_1, \ldots, x_n)^T$. So, for example,

$$(2, 3, 1)^T = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}\,.$$

Let's summarise: the non-empty set L satisfies conditions (i) and (ii), that is, it is closed under addition and scalar multiplication, hence L is a subspace of $\mathbb{R}^3$ as claimed.
(b) In order to see that P is a subspace of $\mathbb{R}^3$ we first note that $(0,0,0)^T \in P$, so P is not empty.

Next we check condition (i). Let $\mathbf{x}_1 = (r_1, s_1, t_1)^T \in P$ and $\mathbf{x}_2 = (r_2, s_2, t_2)^T \in P$. Then $r_1 - s_1 + 3t_1 = 0$ and $r_2 - s_2 + 3t_2 = 0$, so

$$\mathbf{x}_1 + \mathbf{x}_2 = \begin{pmatrix} r_1 + r_2 \\ s_1 + s_2 \\ t_1 + t_2 \end{pmatrix} \in P\,,$$

since $(r_1 + r_2) - (s_1 + s_2) + 3(t_1 + t_2) = (r_1 - s_1 + 3t_1) + (r_2 - s_2 + 3t_2) = 0 + 0 = 0$. Thus condition (i) holds.

We now check condition (ii). Let $\mathbf{x} = (r, s, t)^T \in P$ and let α be a scalar. Then $r - s + 3t = 0$ and

$$\alpha\mathbf{x} = \begin{pmatrix} \alpha r \\ \alpha s \\ \alpha t \end{pmatrix} \in P$$

since $\alpha r - \alpha s + 3\alpha t = \alpha(r - s + 3t) = 0$. Thus condition (ii) holds as well.

As P is closed under addition and scalar multiplication, P is a subspace of $\mathbb{R}^3$ as claimed. □

Remark 4.10. In the example above the two subspaces L and P of $\mathbb{R}^3$ can also be thought of as geometric objects. More precisely, L can be interpreted geometrically as a line through the origin with direction vector $(1,1,1)^T$, while P can be interpreted as a plane through the origin with normal vector $(1,-1,3)^T$.

More generally, all proper subspaces of $\mathbb{R}^3$ can be interpreted geometrically as either lines or planes through the origin. Similarly, all proper subspaces of $\mathbb{R}^2$ can be interpreted geometrically as lines through the origin.

If this is not crystal clear to you, think about it!

Example 4.11. $H = \left\{ (r^2, s, r)^T \,\middle|\, r, s \in \mathbb{R} \right\}$ is not a subspace of $\mathbb{R}^3$, since

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \in H, \text{ but } 2\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix} \notin H\,.$$

Example 4.12. The set

$$H = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2} \,\middle|\, a, b \in \mathbb{R} \right\},$$

is not a subspace of $\mathbb{R}^{2\times 2}$. In order to see this, note that every subspace must contain the zero vector. However,

$$O_{2\times 2} \notin H\,.$$

Example 4.13. Let $H = \{\, \mathbf{f} \in C[-2,2] \mid \mathbf{f}(1) = 0 \,\}$. Then H is a subspace of $C[-2,2]$. First observe that the zero function is in H, so H is not empty. Next we check that the closure properties are satisfied.

Let $\mathbf{f}, \mathbf{g} \in H$. Then $\mathbf{f}(1) = 0$ and $\mathbf{g}(1) = 0$, so

$$(\mathbf{f} + \mathbf{g})(1) = \mathbf{f}(1) + \mathbf{g}(1) = 0 + 0 = 0\,,$$

so $\mathbf{f}+\mathbf{g} \in H$. Thus H is closed under addition.

Let $\mathbf{f} \in H$ and α be a a real number. Then $\mathbf{f}(1) = 0$ and

$$(\alpha\mathbf{f})(1) = \alpha\mathbf{f}(1) = \alpha \cdot 0 = 0\,,$$

so $\alpha\mathbf{f} \in H$. Thus H is closed under scalar multiplication.

Since H is closed under addition and scalar multiplication it is a subspace of $C[-2,2]$ as claimed.

A class of subspaces we have already encountered (but didn't think about them in this way) are the solution sets of homogeneous systems. More precisely, if $A \in \mathbb{R}^{m\times n}$ is the coefficient matrix of such a system, then the solution set can be thought of as the collection of all $\mathbf{x} \in \mathbb{R}^n$ with $A\mathbf{x} = \mathbf{0}$, and the collection of all such $\mathbf{x}$ is a subspace of $\mathbb{R}^n$. Before convincing us of this fact, we introduce some convenient terminology:

Definition 4.14. Let $A \in \mathbb{R}^{m\times n}$. Then

$$N(A) = \{\, \mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0} \,\}$$

is called the **nullspace** of A.

Theorem 4.15. *If $A \in \mathbb{R}^{m\times n}$, then $N(A)$ is a subspace of $\mathbb{R}^n$.*

Proof. Clearly $\mathbf{0} \in N(A)$, so $N(A)$ is not empty.

If $\mathbf{x}, \mathbf{y} \in N(A)$ then $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{y} = \mathbf{0}$, so

$$A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{0} + \mathbf{0} = \mathbf{0}\,,$$

and hence $\mathbf{x}+\mathbf{y} \in N(A)$.

Furthermore, if $\mathbf{x} \in N(A)$ and α is a real number then $A\mathbf{x} = \mathbf{0}$ and

$$A(\alpha\mathbf{x}) = \alpha(A\mathbf{x}) = \alpha\mathbf{0} = \mathbf{0}\,,$$

so $\alpha\mathbf{x} \in N(A)$.

Thus $N(A)$ is a subspace of $\mathbb{R}^n$ as claimed. □

Example 4.16. Determine $N(A)$ for

$$A = \begin{pmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{pmatrix}.$$

Solution. We need to find the solution set of $A\mathbf{x} = \mathbf{0}$. To do this you can use your favourite method to solve linear systems. Perhaps the fastest one is to bring the augmented matrix $(A|\mathbf{0})$ to reduced row echelon form and write the leading variables in terms of the free variables. In our case, we have

$$\left(\begin{array}{ccccc|c} -3 & 6 & -1 & 1 & -7 & 0 \\ 1 & -2 & 2 & 3 & -1 & 0 \\ 2 & -4 & 5 & 8 & -4 & 0 \end{array}\right) \sim \cdots \sim \left(\begin{array}{ccccc|c} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

The leading variables are x_1 and x_3, and the free variables are x_2, x_4 and x_5. Now setting $x_2 = \alpha$, $x_4 = \beta$ and $x_5 = \gamma$ we find $x_3 = -2x_4 + 2x_5 = -2\beta + 2\gamma$ and $x_1 = 2x_2 + x_4 - 3x_5 = 2\alpha + \beta - 3\gamma$. Thus

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 2\alpha + \beta - 3\gamma \\ \alpha \\ -2\beta + 2\gamma \\ \beta \\ \gamma \end{pmatrix} = \alpha \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} + \gamma \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix},$$

hence

$$N(A) = \left\{ \alpha \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} + \gamma \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} \middle| \, \alpha, \beta, \gamma \in \mathbb{R} \right\}.$$

□

4.3 The span of a set of vectors

In this section we shall have a look at a way to construct subspaces from a collection of vectors.

Definition 4.17. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be vectors in a vector space. The set of all linear combinations[4] of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is called the **span** of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and is denoted by $\mathrm{Span}\,(\mathbf{v}_1, \ldots, \mathbf{v}_n)$, that is,

$$\mathrm{Span}\,(\mathbf{v}_1, \ldots, \mathbf{v}_n) = \{\, \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n \mid \alpha_1, \ldots, \alpha_n \in \mathbb{R} \,\}.$$

Example 4.18. Let $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \in \mathbb{R}^3$ be given by

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Determine $\mathrm{Span}\,(\mathbf{e}_1, \mathbf{e}_2)$ and $\mathrm{Span}\,(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$.

Solution. Since

$$\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ 0 \end{pmatrix}, \quad \text{while} \quad \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \alpha_3\mathbf{e}_3 = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix},$$

we see that

$$\mathrm{Span}\,(\mathbf{e}_1, \mathbf{e}_2) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \middle| \, x_3 = 0 \right\}, \quad \text{while} \quad \mathrm{Span}\,(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = \mathbb{R}^3.$$

□

Notice that in the above example $\mathrm{Span}\,(\mathbf{e}_1, \mathbf{e}_2)$ can be interpreted geometrically as the x_1, x_2 plane, that is, the plane containing the x_1- and the x_2-axis. In particular, $\mathrm{Span}\,(\mathbf{e}_1, \mathbf{e}_2)$ is a subspace of $\mathbb{R}^3$. This is true more generally:

[4] In case you forgot what this means go back to Definition 2.30.

Example 4.19. Given vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in a vector space V, show that $H = \mathrm{Span}\,(\mathbf{v}_1, \mathbf{v}_2)$ is a subspace of V.

Solution. Notice that $\mathbf{0} \in H$ (since $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2$), so H is not empty. In order to show that H is closed under addition, let $\mathbf{u}$ and $\mathbf{w}$ be arbitrary vectors in H. Then there are scalars α_1, α_2 and β_1, β_2, such that

$$\mathbf{u} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2\,,$$
$$\mathbf{w} = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2\,.$$

Now by axioms (A1), (A2) and (A6)

$$\mathbf{u} + \mathbf{w} = (\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) + (\beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2) = (\alpha_1 + \beta_1)\mathbf{v}_1 + (\alpha_2 + \beta_2)\mathbf{v}_2\,,$$

so $\mathbf{u} + \mathbf{w} \in H$.

In order to show that H is closed under scalar multiplication, let $\mathbf{u} \in H$, say, $\mathbf{u} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2$, and let γ be a scalar. Then, by axioms (A5) and (A7)

$$\gamma\mathbf{u} = \gamma(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = (\gamma\alpha_1)\mathbf{v}_1 + (\gamma\alpha_2)\mathbf{v}_2\,,$$

so $\gamma\mathbf{u} \in H$. □

More generally, using exactly the same method of proof, it is possible to show the following:

Theorem 4.20. *Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be vectors in a vector space V. Then $\mathrm{Span}\,(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ is a subspace of V.*

We have just seen that the span of a collection of vectors in a vector space V is a subspace of V. As we saw in Example 4.18, the span may be a proper subspace of V, or it may be equal to all of V. The latter is sufficiently interesting a case to merit its own definition:

Definition 4.21. Let V be a vector space, and let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$. We say that the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a **spanning set** for V if

$$\mathrm{Span}\,(\mathbf{v}_1, \ldots, \mathbf{v}_n) = V\,.$$

If $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a spanning set for V, we shall also say that $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ **spans** V, that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ **span** V or that V is **spanned** by $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

Notice that the above definition can be rephrased as follows. A set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a spanning set for V, if and only if every vector in V can be written as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

Example 4.22. Which of the following sets are spanning sets for $\mathbb{R}^3$?

(a) $\left\{(1,0,0)^T, (0,1,0)^T, (0,0,1)^T, (1,2,4)^T\right\}$ (b) $\left\{(1,1,1)^T, (1,1,0)^T, (1,0,0)^T\right\}$

(c) $\left\{(1,0,1)^T, (0,1,0)^T\right\}$ (d) $\left\{(1,2,4)^T, (2,1,3)^T, (4,-1,1)^T\right\}$

Solution. (a) Let $(a,b,c)^T$ be an arbitrary vector in $\mathbb{R}^3$. Clearly

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}\,,$$

so the set is a spanning set for $\mathbb{R}^3$.
(b) Let $(a, b, c)^T$ be an arbitrary vector in $\mathbb{R}^3$. We need to determine whether it is possible to find constants α_1, α_2, α_3 such that

$$\alpha_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} .$$

This means that the weights α_1, α_2 and α_3 have to satisfy the system

$$\begin{array}{rcrcrcl} \alpha_1 & + & \alpha_2 & + & \alpha_3 & = & a \\ \alpha_1 & + & \alpha_2 & & & = & b \\ \alpha_1 & & & & & = & c \end{array}$$

Since the coefficient matrix of the system is nonsingular, the system has a unique solution. In fact, using back substitution we find

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} c \\ b - c \\ a - b \end{pmatrix} .$$

Thus

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = c \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + (b - c) \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + (a - b) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} ,$$

so the set is a spanning set for $\mathbb{R}^3$.
(c) Noting that

$$\alpha_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_1 \end{pmatrix} ,$$

we see that a vector of the form $(a, b, c)^T$ with $a \neq c$ cannot be in the span of the two vectors. Thus the set is not a spanning set for $\mathbb{R}^3$.
(d) Proceeding as in (b), we let $(a, b, c)^T$ be an arbitrary vector in $\mathbb{R}^3$. Again, we need to determine whether it is possible to find constants α_1, α_2, α_3 such that

$$\alpha_1 \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} + \alpha_2 \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + \alpha_3 \begin{pmatrix} 4 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} .$$

This means that the weights α_1, α_2 and α_3 have to satisfy the system

$$\begin{array}{rcrcrcl} \alpha_1 & + & 2\alpha_2 & + & 4\alpha_3 & = & a \\ 2\alpha_1 & + & \alpha_2 & - & \alpha_3 & = & b \\ 4\alpha_1 & + & 3\alpha_2 & + & \alpha_3 & = & c \end{array}$$

A short calculation shows that the coefficient matrix of the system is singular, from which we could conclude that the system cannot have a solution for all $a, b, c \in \mathbb{R}$. In other words, the vectors cannot span $\mathbb{R}^3$. It is however instructive to reach the same conclusion by a slightly different route: using Gaussian elimination we see that the system is equivalent to the following

$$\begin{array}{rcrcrcl} \alpha_1 & + & 2\alpha_2 & + & 4\alpha_3 & = & a \\ & & \alpha_2 & + & 3\alpha_3 & = & \frac{2a-b}{3} \\ & & & & 0 & = & 2a + 5b - 3c \end{array}$$

It follows that the system is consistent if and only if

$$2a + 5b - 3c = 0\,.$$

Thus a vector $(a, b, c)^T$ in $\mathbb{R}^3$ belongs to the span of the vectors $(1, 2, 4)^T$, $(2, 1, 3)^T$, and $(4, -1, 1)^T$ if and only if $2a + 5b - 3c = 0$. In other words, not every vector in $\mathbb{R}^3$ can be written as a linear combination of the vectors $(1, 2, 4)^T$, $(2, 1, 3)^T$, and $(4, -1, 1)^T$, so in particular these vectors cannot span $\mathbb{R}^3$. □

Example 4.23. Show that $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ is a spanning set for P_2, where

$$\mathbf{p}_1(x) = 2 + 3x + x^2\,, \quad \mathbf{p}_2(x) = 4 - x\,, \quad \mathbf{p}_3(x) = -1\,.$$

Solution. Let $\mathbf{p}$ be an arbitrary polynomial in P_2, say, $\mathbf{p}(x) = a + bx + cx^2$. We need to show that it is possible to find weights α_1, α_2 and α_3 such that

$$\alpha_1\mathbf{p}_1 + \alpha_2\mathbf{p}_2 + \alpha_3\mathbf{p}_3 = \mathbf{p}\,,$$

that is

$$\alpha_1(2 + 3x + x^2) + \alpha_2(4 - x) - \alpha_3 = a + bx + cx^2\,.$$

Comparing coefficients we find that the weights have to satisfy the system

$$\begin{array}{rcrcrcl} 2\alpha_1 & + & 4\alpha_2 & - & \alpha_3 & = & a \\ 3\alpha_1 & - & \alpha_2 & & & = & b \\ \alpha_1 & & & & & = & c \end{array}$$

The coefficient matrix is nonsingular, so the system must have a unique solution for all choices of a, b, c. In fact, using back substitution yields $\alpha_1 = c$, $\alpha_2 = 3c - b$, $\alpha_3 = 14c - 4b - a$. Thus $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ is a spanning set for P_2. □

Example 4.24. Find a spanning set for $N(A)$, where

$$A = \begin{pmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{pmatrix}\,.$$

Proof. We have already calculated $N(A)$ for this matrix in Example 4.16, and found that

$$N(A) = \left\{ \alpha \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} + \gamma \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} \,\middle|\, \alpha, \beta, \gamma \in \mathbb{R} \right\}\,.$$

Thus, $\left\{(2, 1, 0, 0, 0)^T, (1, 0, -2, 1, 0)^T, (-3, 0, 2, 0, 1)^T\right\}$ is a spanning set for $N(A)$. □

4.4 Linear independence

The notion of linear independence plays a fundamental role in the theory of vector spaces. Roughly speaking, it is a certain minimality property a collection of vectors in a vector space may or may not have. To motivate it consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{x}_1 = \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix}\,, \quad \mathbf{x}_2 = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}\,, \quad \mathbf{x}_3 = \begin{pmatrix} -5 \\ 1 \\ 8 \end{pmatrix}\,. \tag{4.3}$$

Let's ask the question: what is $\mathrm{Span}\,(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$?

Notice that

$$\mathbf{x}_3 = 3\mathbf{x}_1 + 2\mathbf{x}_2\,. \tag{4.4}$$

Thus any linear combination of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ can be written as a linear combination of $\mathbf{x}_1$ and $\mathbf{x}_2$ alone, because

$$\begin{aligned}\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \alpha_3\mathbf{x}_3 &= \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \alpha_3(3\mathbf{x}_1 + 2\mathbf{x}_2)\\ &= (\alpha_1 + 3\alpha_2)\mathbf{x}_1 + (\alpha_2 + 2\alpha_3)\mathbf{x}_2\,.\end{aligned}$$

Hence

$$\mathrm{Span}\,(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \mathrm{Span}\,(\mathbf{x}_1, \mathbf{x}_2)\,.$$

Observing that equation (4.4) can be written as

$$3\mathbf{x}_1 + 2\mathbf{x}_2 - \mathbf{x}_3 = \mathbf{0}\,, \tag{4.5}$$

we see that any of the three vectors can be expressed as a linear combination of the other two, so

$$\mathrm{Span}\,(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \mathrm{Span}\,(\mathbf{x}_1, \mathbf{x}_2) = \mathrm{Span}\,(\mathbf{x}_1, \mathbf{x}_3) = \mathrm{Span}\,(\mathbf{x}_2, \mathbf{x}_3)\,.$$

In other words, because of the dependence relation (4.5), the span of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ can be written as the span of only two of the given vectors. Or, put yet differently, we can throw away one of the three vectors without changing their span. So the three vectors are not the most economic way to express their span, because two of them suffice.

On the other hand, no dependency of the form (4.5) exists between $\mathbf{x}_1$ and $\mathbf{x}_2$. Indeed, suppose that there were scalars c_1 and c_2, not both 0, such that

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 = \mathbf{0}\,, \tag{4.6}$$

then we could solve for one of the vectors in terms of the other, that is,

$$\mathbf{x}_1 = -\frac{c_2}{c_1}\mathbf{x}_2 \quad (\text{if } c_1 \neq 0) \qquad \text{or} \qquad \mathbf{x}_2 = -\frac{c_1}{c_2}\mathbf{x}_1 \quad (\text{if } c_2 \neq 0)\,.$$

However, neither vector is a multiple of the other, so (4.6) can hold only if $c_1 = c_2 = 0$. In particular, $\mathrm{Span}\,(\mathbf{x}_1)$ and $\mathrm{Span}\,(\mathbf{x}_2)$ are both proper subspaces of $\mathrm{Span}\,(\mathbf{x}_1, \mathbf{x}_2)$, so we cannot reduce the number of vectors further to express $\mathrm{Span}\,(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \mathrm{Span}\,(\mathbf{x}_1, \mathbf{x}_2)$.

This discussion motivates the following definitions:

Definition 4.25. The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in a vector space V are said to be **linearly dependent** if there exist scalars $c_1, \ldots, c_n$, not all zero, such that

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}\,.$$

Example 4.26. The three vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ defined in (4.3) are linearly dependent.

Definition 4.27. The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in a vector space V are said to be **linearly independent** if they are not linearly dependent, that is, if

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}\,,$$

forces all scalars $c_1, \ldots, c_n$ to be 0.

Example 4.28. The vectors $\begin{pmatrix}2\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix} \in \mathbb{R}^2$ are linearly independent. In order to see this, suppose that

$$c_1\begin{pmatrix}2\\1\end{pmatrix} + c_2\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}.$$

Then c_1 and c_2 must satisfy the 2×2 system

$$\begin{array}{rcrcl} 2c_1 & + & c_2 & = & 0 \\ c_1 & + & c_2 & = & 0 \end{array}$$

However, as is easily seen, the only solution of this system is $c_1 = c_2 = 0$. Thus, the two vectors are indeed linearly independent as claimed.

Example 4.29. Let $\mathbf{p}_1, \mathbf{p}_2 \in P_1$ be given by

$$\mathbf{p}_1(t) = 2 + t\,, \quad \mathbf{p}_2(t) = 1 + t\,.$$

Then $\mathbf{p}_1$ and $\mathbf{p}_2$ are linearly independent. In order to see this, suppose that

$$c_1\mathbf{p}_1 + c_2\mathbf{p}_2 = \mathbf{0}\,.$$

Then, for all t

$$c_1(2 + t) + c_2(1 + t) = 0\,,$$

so, for all t

$$(2c_1 + c_2) + (c_1 + c_2)t = 0\,.$$

Notice that the polynomial on the left-hand side of the above equation will be the zero polynomial if and only if its coefficients vanish, so c_1 and c_2 must satisfy the 2×2 system

$$\begin{array}{rcrcl} 2c_1 & + & c_2 & = & 0 \\ c_1 & + & c_2 & = & 0 \end{array}$$

However, as in the previous example, the only solution of this system is $c_1 = c_2 = 0$. Thus $\mathbf{p}_1$ and $\mathbf{p}_2$ are indeed linearly independent as claimed.

Example 4.30 (Geometric interpretation of linear independence in $\mathbb{R}^2$ and $\mathbb{R}^3$)**.** In case you missed the hands-on demonstration with audience participation of linear independence given in the lectures, make sure you hard-wire the following mental images of linear independence in $\mathbb{R}^2$ and $\mathbb{R}^3$ to your brain for the rest of this module:

(a) If $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent in $\mathbb{R}^2$ then

$$c_1\mathbf{x} + c_2\mathbf{y} = \mathbf{0}\,,$$

where c_1 and c_2 are not both 0. If, say $c_1 \neq 0$, then

$$\mathbf{x} = -\frac{c_2}{c_1}\mathbf{y}\,.$$

Thus one of the vectors must be a scalar multiple of the other, or, put differently, the two vectors must be collinear.

Conversely, if two vectors in $\mathbb{R}^2$ are not collinear, they are linearly independent.

(b) Just as in $\mathbb{R}^2$, two vectors in $\mathbb{R}^3$ are linearly dependent if and only if they are collinear. Suppose now that $\mathbf{x}$ and $\mathbf{y}$ are two linearly independent vectors in $\mathbb{R}^3$. Since they are not collinear, they will span a plane (through the origin). If $\mathbf{z}$ is another vector lying in this plane, then $\mathbf{0}$ can be written as a linear combination of $\mathbf{x}$ and $\mathbf{y}$, hence $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are linearly dependent. Conversely, if $\mathbf{z}$ does not lie in the plane spanned by $\mathbf{x}$ and $\mathbf{y}$, then $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are linearly independent.

In other words, three vectors in $\mathbb{R}^3$ are linearly independent if and only if they are not coplanar.

We finish this section with some more theoretical observations:

Theorem 4.31. *Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be n vectors in $\mathbb{R}^n$ and let $X \in \mathbb{R}^{n \times n}$ be the matrix whose j-th column is $\mathbf{x}_j$. Then the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are linearly dependent if and only if X is singular.*

Proof. The equation

$$c_1\mathbf{x}_1 + \cdots + c_n\mathbf{x}_n = \mathbf{0}$$

can be written as

$$X\mathbf{c} = \mathbf{0}\,, \quad \text{where} \quad \mathbf{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}\,.$$

This system has a non-trivial solution $\mathbf{c} \neq \mathbf{0}$ if and only X is singular. □

Example 4.32. Determine whether the following three vectors in $\mathbb{R}^3$ are linearly independent:

$$\begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix}\,, \quad \begin{pmatrix} 5 \\ 2 \\ 5 \end{pmatrix}\,, \quad \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}\,.$$

Solution. Since

$$\begin{vmatrix} -1 & 5 & 4 \\ 3 & 2 & 5 \\ 1 & 5 & 6 \end{vmatrix} = \begin{vmatrix} -1 & 3 & 1 \\ 5 & 2 & 5 \\ 4 & 5 & 6 \end{vmatrix} \begin{matrix} R_1 + R_2 \\ = \end{matrix} \begin{vmatrix} 4 & 5 & 6 \\ 5 & 2 & 5 \\ 4 & 5 & 6 \end{vmatrix} \begin{matrix} R_1 - R_3 \\ = \end{matrix} \begin{vmatrix} 0 & 0 & 0 \\ 5 & 2 & 5 \\ 4 & 5 & 6 \end{vmatrix} = 0\,,$$

the vectors are linearly dependent. □

The following result will become important later in this chapter, when we discuss coordinate systems.

Theorem 4.33. *Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be vectors in a vector space V. A vector $\mathbf{v} \in \operatorname{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ can be written uniquely as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ if and only if $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent.*

Proof. If $\mathbf{v} \in \operatorname{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ then $\mathbf{v}$ can be written

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n\,, \tag{4.7}$$

for some scalars $\alpha_1, \ldots, \alpha_n$. Suppose that $\mathbf{v}$ can also be written in the form

$$\mathbf{v} = \beta_1\mathbf{v}_1 + \cdots + \beta_n\mathbf{v}_n\,, \tag{4.8}$$

for some scalars $\beta_1, \ldots, \beta_n$. We start by showing that if $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent, then $\alpha_i = \beta_i$ for every $i = 1, \ldots, n$ (that is, the representation (4.7) is unique). To see this, suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent. Then subtracting (4.8) from (4.7) gives

$$(\alpha_1 - \beta_1)\mathbf{v}_1 + \cdots + (\alpha_n - \beta_n)\mathbf{v}_n = \mathbf{0}\,, \tag{4.9}$$

which forces $\alpha_i = \beta_i$ for every $i = 1, \ldots, n$ as desired.

Conversely, if the representation (4.7) is not unique, then there must be a representation of the form (4.8) where $\alpha_i \neq \beta_i$ for some i between 1 and n. But then (4.9) means that there exists a non-trivial linear dependence between $\mathbf{v}_1, \ldots, \mathbf{v}_n$, so these vectors are linearly dependent. □

4.5 Basis and dimension

The concept of a basis and the related notion of dimension are among the key ideas in the theory vector of spaces, of immense practical and theoretical importance. Let's start with the definition of a basis, delaying the discussion of its interpretation for a bit:

Definition 4.34. A set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of vectors forms a **basis** for a vector space V if

(i) $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent;

(ii) $\mathrm{Span}\,(\mathbf{v}_1, \ldots, \mathbf{v}_n) = V$.

In other words, a basis for a vector space is a 'minimal' spanning set, in the sense that it contains no superfluous vectors: every vector in V can be written as a linear combination of the basis vectors (because of property (ii)), and there is no redundancy in the sense that no basis vector can be expressed as a linear combination of the other basis vectors (by property (i)). Let's look at some examples:

Example 4.35. Let

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\,, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\,, \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\,.$$

Then $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is a basis for $\mathbb{R}^3$, called the **standard basis**.

Indeed, as is easily seen, every vector in $\mathbb{R}^3$ can be written as a linear combination of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ and, moreover, the vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are linearly independent.

Example 4.36.

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

is a basis for $\mathbb{R}^3$.

To see this note that the vectors are linearly independent, because

$$\begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} = 1 \neq 0\,.$$

Moreover, the vectors span $\mathbb{R}^3$ since, if $(a, b, c)^T$ is an arbitrary vector in $\mathbb{R}^3$, then

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = (a - b) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (b - c) \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + c \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\,.$$

The previous two examples show that a vector space may have more than one basis. This is not a nuisance, but, quite to the contrary, a blessing, as we shall see later in this module. For the moment, you should only note that both bases consist of exactly three elements. We will revisit and expand this seemingly innocent observation shortly, when we discuss the dimension of a vector space.

Example 4.37. Let

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then $\{E_{11}, E_{12}, E_{21}, E_{22}\}$ is a basis for $\mathbb{R}^{2\times 2}$, because the four vectors span $\mathbb{R}^{2\times 2}$ (as was shown in Coursework 5, Exercise 7(b)) and they are linearly independent. To see this, suppose that

$$c_1E_{11} + c_2E_{12} + c_3E_{21} + c_4E_{22} = O_{2\times 2}\,.$$

Then

$$\begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

so $c_1 = c_2 = c_3 = c_4 = 0$.

Most of the vector spaces we have encountered so far have particularly simple bases, termed 'standard bases':

Example 4.38 (Standard bases for $\mathbb{R}^n$, $\mathbb{R}^{m\times n}$ and P_n)**.**

$\mathbb{R}^n$: The n columns of I_n form the standard basis of $\mathbb{R}^n$, usually denoted by $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$.

$\mathbb{R}^{m\times n}$: A canonical basis can be constructed as follows. For $i = 1, \ldots, m$ and $j = 1, \ldots, n$ let $E_{ij} \in \mathbb{R}^{m\times n}$ be the matrix whose (i, j)-entry is 1, and all other entries are 0. Then $\{\, E_{ij} \mid i = 1, \ldots, m\,, j = 1, \ldots, n\,\}$ is the standard basis for $\mathbb{R}^{m\times n}$.

P_n: The standard basis is the collection $\{\mathbf{p}_0, \ldots, \mathbf{p}_n\}$ of all monomials of degree less than n, that is,

$$\mathbf{p}_k(t) = t^k\,, \quad \text{for } k = 0, \ldots, n.$$

If this is not clear to you, you should check that it really is a basis!

Going back to Examples 4.35 and 4.36, recall the observation that both bases of $\mathbb{R}^3$ contained exactly three elements. This is not pure coincidence, but has a deeper reason. In fact, as we shall see shortly, any basis of a vector space must contain the same number of vectors. Before proving this important result we need the following:

Theorem 4.39. *Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be vectors in a vector space V. If $\operatorname{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n) = V$, then any collection of m vectors in V where $m > n$ is linearly dependent.*

Proof. Let $\mathbf{u}_1, \ldots, \mathbf{u}_m$ be m vectors in V where $m > n$. Then, since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ span V, we can write

$$\mathbf{u}_i = \alpha_{i1}\mathbf{v}_1 + \cdots + \alpha_{in}\mathbf{v}_n \quad \text{for } i = 1, \ldots, m\,.$$

Thus, a linear combination of the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_m$ can be written as

$$\begin{aligned} c_1\mathbf{u}_1 + \cdots + c_m\mathbf{u}_m &= \sum_{i=1}^{m} c_i\mathbf{u}_i \\ &= \sum_{i=1}^{m} c_i \left(\sum_{j=1}^{n} \alpha_{ij}\mathbf{v}_j \right) \\ &= \sum_{j=1}^{n} \left(\sum_{i=1}^{m} \alpha_{ij}c_i \right) \mathbf{v}_j\,. \end{aligned}$$

Now consider the system of n equations for the m unknowns $c_1, \ldots, c_m$

$$\sum_{i=1}^{m} \alpha_{ij} c_i = 0 \quad \text{for } j = 1, \ldots, n.$$

This is a homogeneous system with more unknowns than equations, so by Theorem 1.22 it must have a non-trivial solution $(\hat{c}_1, \ldots, \hat{c}_m)^T$. But then

$$\hat{c}_1 \mathbf{u}_1 + \cdots + \hat{c}_m \mathbf{u}_m = \sum_{j=1}^{n} 0 \mathbf{v}_j = \mathbf{0}\,,$$

so $\mathbf{u}_1, \ldots, \mathbf{u}_m$ are linearly dependent. □

We are now able to prove the observation alluded to earlier:

Corollary 4.40. *If a vector space V has a basis of n vectors, then every basis of V must have exactly n vectors.*

Proof. Suppose that $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ are both bases for V. We shall show that $m = n$. In order to see this, notice that, since $\mathrm{Span}\,(\mathbf{v}_1, \ldots, \mathbf{v}_n) = V$ and $\mathbf{u}_1, \ldots, \mathbf{u}_m$ are linearly independent it follows by the previous theorem that $m \leq n$. By the same reasoning, since $\mathrm{Span}\,(\mathbf{u}_1, \ldots, \mathbf{u}_m) = V$ and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent, we must have $n \leq m$. So, all in all, we have $n = m$, that is, the two bases have the same number of elements. □

In view of this corollary it now makes sense to talk about *the* number of elements of a basis, and give it a special name:

Definition 4.41. Let V be a vector space. If V has a basis consisting of n vectors, we say that V has **dimension** n, and write $\dim V = n$.

The vector space $\{\mathbf{0}\}$ is said to have dimension 0. The vector space V is said to be **finite dimensional** if there is a finite set of vectors spanning V; otherwise it is said to be **infinite dimensional** .

Example 4.42. By Example 4.38 the vector spaces $\mathbb{R}^n$, $\mathbb{R}^{m\times n}$ and P_n are finite dimensional with dimensions

$$\dim \mathbb{R}^n = n\,, \quad \dim \mathbb{R}^{m\times n} = mn\,, \quad \dim P_n = n+1\,.$$

As an example of an infinite dimensional vector space, consider the set of all polynomials with real coefficients, and call it P. In order to see that P is a vector space when equipped with the usual addition and scalar multiplication, notice that P is a subset of the vector space $C[-1,1]$ of continuous functions on $[-1,1]$ (in fact, it is a subset of $C[a,b]$ for any $a, b \in \mathbb{R}$ with $a < b$), which is closed under addition and scalar multiplication. Thus P is a vector space. Note that any finite collection of monomials is linearly independent, so P must be infinite dimensional. For the same reason, $C[a,b]$ and $C^1[a,b]$ are infinite dimensional vector spaces. While infinite dimensional vector spaces play an important role in many parts of contemporary applied and pure mathematics, we shall be mainly concerned with finite dimensional vector spaces for the rest of this module.

Example 4.43. We are now able to classify the subspaces of $\mathbb{R}^3$ discussed in Remark 4.10:

- 0-*dimensional subspaces*. Only the zero subspace $\{\mathbf{0}\}$.

- 1-*dimensional subspaces*. Any subspace spanned by a nonzero vector, that is all lines through the origin.
- 2-*dimensional subspaces*. Any subspace spanned by two linearly independent vectors, that is all planes through the origin.
- 3-*dimensional subspaces*. Only $\mathbb{R}^3$.

We close this section with the following result, which is often useful when trying to decide whether a collection of vectors forms a basis of a vector space:

Theorem 4.44. *If V is a vector space with $\dim V = n$, then:*

(a) any set consisting of n linearly independent vectors spans V;

(b) any n vectors that span V are linearly independent.

Proof. (a) Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a linearly independent. Pick $\mathbf{v} \in V$. Since $\dim V = n$, the $n+1$ vectors $\mathbf{v}, \mathbf{v}_1, \ldots, \mathbf{v}_n$ must be linearly dependent by Theorem 4.39. Thus

$$c_0\mathbf{v} + c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}\,, \tag{4.10}$$

where $c_0, c_1, \ldots, c_n$ are not all 0. But $c_0 \neq 0$ (for otherwise (4.10) would imply that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly dependent), hence

$$\mathbf{v} = \left(-\frac{c_1}{c_0}\right)\mathbf{v}_1 + \cdots + \left(-\frac{c_n}{c_0}\right)\mathbf{v}_n\,,$$

so $\mathbf{v} \in \operatorname{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$. But $\mathbf{v}$ was arbitrary, so $\operatorname{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n) = V$.

(b) Suppose $\operatorname{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n) = V$. In order to show that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent we argue by contradiction: suppose to the contrary that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly dependent. Then one of the $\mathbf{v}_i$'s, say $\mathbf{v}_n$ can be written as a linear combination of the other vectors. So $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ also span V. If $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ are still linearly dependent we can eliminate another vector and still have a spanning set. We can continue this process until we have found a spanning set with k linearly independent vectors where $k < n$. This, however, contradicts the fact that $\dim V = n$. □

Remark 4.45. The above theorem provides a convenient tool to check whether a set of vectors forms a basis. The theorem tells us that n linearly independent vectors in an n-dimensional vector space are automatically spanning, so these vectors are a basis for the vector space. This is often useful in situations where linear independence is easier to check than the spanning property.

Remark 4.46. The above theorem also provides two perspectives on a basis of a vector space:

$$\text{a basis is } \begin{cases} \text{a spanning set that is as small as possible;} \\ \text{a linearly independent collection of vectors that is as large as possible.} \end{cases}$$

So, for example:

$$\underbrace{\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}2\\3\\0\end{pmatrix}\right\}}_{\substack{\text{linearly independent,}\\ \text{but doesn't span } \mathbb{R}^3}}, \quad \underbrace{\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}2\\3\\0\end{pmatrix}, \begin{pmatrix}4\\5\\6\end{pmatrix}\right\}}_{\text{basis for } \mathbb{R}^3}, \quad \underbrace{\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}2\\3\\0\end{pmatrix}, \begin{pmatrix}4\\5\\6\end{pmatrix}, \begin{pmatrix}7\\8\\9\end{pmatrix}\right\}}_{\substack{\text{spans } \mathbb{R}^3,\\ \text{but not linearly independent}}}.$$

4.6 Coordinates

In this short section we shall discuss an important application of the notion of a basis. In essence, a basis allows us to view a vector space of dimension n as if it were $\mathbb{R}^n$. This is a tremendously useful idea, with many practical and theoretical applications, many of which you will see in the following chapters.

The basic idea is the following. Suppose that $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is a basis for a vector space V. Since the basis vectors are spanning, given $\mathbf{v} \in V$, there are scalars $c_1, \ldots, c_n$ such that

$$\mathbf{v} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n\,.$$

Moreover, since the basis vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n$ are linearly independent, the scalars $c_1, \ldots, c_n$ are uniquely determined by Theorem 4.33. Thus, the vector $\mathbf{v}$ in the vector space V, can be uniquely represented as an n-vector $(c_1, \ldots, c_n)^T$ in $\mathbb{R}^n$. This motivates the following definition:

Definition 4.47. Suppose that $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is a basis for a vector space V. If $\mathbf{v} \in V$ then the uniquely determined scalars $c_1, \ldots, c_n$ such that

$$\mathbf{v} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n\,,$$

are called the **coordinates of $\mathbf{v}$ relative to $\mathcal{B}$** . The n-vector $(c_1, \ldots, c_n)^T \in \mathbb{R}^n$ is called the **$\mathcal{B}$-coordinate vector of $\mathbf{v}$**, or the **coordinate vector of $\mathbf{v}$ relative to $\mathcal{B}$**, and is denoted by $[\mathbf{v}]_{\mathcal{B}}$.

Example 4.48. Consider the basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ for $\mathbb{R}^2$, where

$$\mathbf{b}_1 = \begin{pmatrix}1\\0\end{pmatrix}\,, \quad \mathbf{b}_2 = \begin{pmatrix}1\\2\end{pmatrix}\,.$$

Suppose that $\mathbf{x} \in \mathbb{R}^2$ has $\mathcal{B}$-coordinate vector $[\mathbf{x}]_{\mathcal{B}} = (-2, 3)^T$. Find $\mathbf{x}$.

Solution.

$$\mathbf{x} = -2\mathbf{b}_1 + 3\mathbf{b}_2 = (-2)\begin{pmatrix}1\\0\end{pmatrix} + 3\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}1\\6\end{pmatrix}\,.$$

□

Example 4.49. The entries of $\mathbf{x} = \begin{pmatrix}1\\6\end{pmatrix}$ are the coordinates of $\mathbf{x}$ relative to the standard basis $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$, since

$$\begin{pmatrix}1\\6\end{pmatrix} = 1\begin{pmatrix}1\\0\end{pmatrix} + 6\begin{pmatrix}0\\1\end{pmatrix} = 1\mathbf{e}_1 + 6\mathbf{e}_2\,.$$

Thus, $\mathbf{x} = [\mathbf{x}]_{\mathcal{E}}$.

Remark 4.50. If you missed the lecture where I gave a graphical demonstration of the representation of the vector $\begin{pmatrix}1\\6\end{pmatrix} \in \mathbb{R}^2$ with respect to the two bases $\mathcal{B}$ and $\mathcal{E}$ given in the previous two examples, you should spend two minutes pondering these two representations. I'll put two pictures up at this spot, once I have figured out how to do this.

The following problems arise naturally in this context:

- Given a basis $\mathcal{B}$ for a vector space V and the $\mathcal{B}$-coordinate vector $[\mathbf{v}]_{\mathcal{B}}$ for some $\mathbf{v} \in V$, find $\mathbf{v}$.[5]
- Given a basis $\mathcal{B}$ for a vector space V and a vector $\mathbf{v} \in V$, find the $\mathcal{B}$-coordinate vector $[\mathbf{v}]_{\mathcal{B}}$.
- Given two bases $\mathcal{B}$ and $\mathcal{D}$ for a vector space V and $[\mathbf{v}]_{\mathcal{D}}$, find $[\mathbf{v}]_{\mathcal{B}}$.

We shall now address these problems in the case where V is $\mathbb{R}^n$, deferring the general discussion to the next chapter, when we will have better machinery to deal with these problems.

Theorem 4.51. *Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be a basis for $\mathbb{R}^n$. There is an invertible $n \times n$ matrix $P_{\mathcal{B}}$ such that for any $\mathbf{x} \in \mathbb{R}^n$*

$$\mathbf{x} = P_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}\,.$$

In fact, the matrix $P_{\mathcal{B}}$ is the matrix whose j-th column is $\mathbf{b}_j$.

Proof. Let $[\mathbf{x}]_{\mathcal{B}} = (c_1, \ldots, c_n)^T$. Then

$$\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n\,,$$

so

$$\mathbf{x} = \underbrace{(\mathbf{b}_1 \ \cdots \ \mathbf{b}_n)}_{=P_{\mathcal{B}}} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}\,.$$

Moreover, by Theorem 4.31, the matrix $P_{\mathcal{B}}$ is invertible since its columns are linearly independent. □

Since a vector $\mathbf{x} \in \mathbb{R}^n$ is equal to its coordinate vector relative to the standard basis, the matrix $P_{\mathcal{B}}$ given in the theorem above is called the **transition matrix from $\mathcal{B}$ to the standard basis**.

Corollary 4.52. *Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be a basis for $\mathbb{R}^n$. For any $\mathbf{x} \in \mathbb{R}^n$*

$$[\mathbf{x}]_{\mathcal{B}} = P_{\mathcal{B}}^{-1}\mathbf{x}\,.$$

Example 4.53. Let $\mathbf{b}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $\mathbf{b}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$. Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ be the corresponding basis for $\mathbb{R}^2$. Find the $\mathcal{B}$-coordinates of $\mathbf{x}$.

Solution. By the previous corollary

$$[\mathbf{x}]_{\mathcal{B}} = P_{\mathcal{B}}^{-1}\mathbf{x}\,.$$

Now

$$P_{\mathcal{B}} = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}\,,$$

so

$$P_{\mathcal{B}}^{-1} = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}\,.$$

Thus

$$[\mathbf{x}]_{\mathcal{B}} = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}\,.$$

□

[5] We have already done a particular case in Example 4.48.

Theorem 4.54. *Let $\mathcal{B}$ and $\mathcal{D}$ be two bases for $\mathbb{R}^n$. If $\mathbf{x}$ is any vector in $\mathbb{R}^n$, then*

$$[\mathbf{x}]_{\mathcal{B}} = P_{\mathcal{B}}^{-1} P_{\mathcal{D}} [\mathbf{x}]_{\mathcal{D}} .$$

Proof. Clear, since $P_{\mathcal{B}}$ is invertible and

$$P_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}} = \mathbf{x} = P_{\mathcal{D}}[\mathbf{x}]_{\mathcal{D}} .$$

□

The $n \times n$ matrix $P_{\mathcal{B}}^{-1} P_{\mathcal{D}}$ given in the theorem above is called the **transition matrix from $\mathcal{D}$ to $\mathcal{B}$**.

Example 4.55. Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ be the basis given in Example 4.53, let $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$, where

$$\mathbf{d}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \quad \mathbf{d}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} ,$$

and let $\mathbf{x} \in \mathbb{R}^2$. If the $\mathcal{D}$-coordinates of $\mathbf{x}$ are $(-3, 2)^T$, what are the $\mathcal{B}$-coordinates of $\mathbf{x}$?

Solution.

$$[\mathbf{x}]_{\mathcal{B}} = P_{\mathcal{B}}^{-1} P_{\mathcal{D}} [\mathbf{x}]_{\mathcal{D}} = \frac{1}{3} \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} -3 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} .$$

□

4.7 Row space and column space

In this final section of this rather long chapter on vector spaces, we shall briefly discuss a number of naturally arising vector spaces associated with matrices. We have already encountered one such space, the nullspace of a matrix.

Definition 4.56. Let $A \in \mathbb{R}^{m \times n}$.

- The subspace of $\mathbb{R}^{1 \times n}$ spanned by the row vectors of A is called the **row space** of A and is denoted by $\mathrm{row}(A)$.
- The subspace of $\mathbb{R}^{m \times 1}$ spanned by the column vectors of A is called the **column space** of A and is denoted by $\mathrm{col}(A)$.

Example 4.57. Let $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

- Since
$$\alpha \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} + \beta \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} \alpha & \beta & 0 \end{pmatrix}$$
$\mathrm{row}(A)$ is a 2-dimensional subspace of $\mathbb{R}^{1 \times 3}$.
- Since
$$\alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$
$\mathrm{col}(A)$ is a 2-dimensional subspace of $\mathbb{R}^{2 \times 1}$.

Notice that the row space and column space of a matrix are generally distinct objects. Indeed, one is a subspace of $\mathbb{R}^{1\times n}$ the other a subspace of $\mathbb{R}^{m\times 1}$. However, in the example above, both spaces have the same dimension (namely 2). We shall see shortly, that, rather surprisingly, this is always the case. Before exploring this topic further we introduce the following important concept:

Definition 4.58. The **rank** of a matrix, denoted by $\operatorname{rank} A$, is the dimension of the row space.

How does one calculate the rank of a matrix? The next result provides the clue:

Theorem 4.59. *Two row equivalent matrices have the same row space, so, in particular, have the same rank.*

Proof. Let A and B be two row equivalent matrices. Since B is row equivalent to A, the matrix B can be obtained from A by a finite sequence of elementary row operations. Thus the rows of B are a linear combination of the rows of A. Consequently, $\operatorname{row}(B)$ is a subspace of $\operatorname{row}(A)$. Exchanging the roles of A and B it follows, using the same argument, that $\operatorname{row}(A)$ is also a subspace of $\operatorname{row}(B)$, so $\operatorname{row}(A) = \operatorname{row}(B)$. □

Combining the previous theorem with the observation that the nonzero rows of a matrix in row echelon form are linearly independent, we obtain the following recipe for calculating a basis for the row space and the rank of a matrix:

In order to calculate a basis for the row space and the rank of a matrix A:

- bring matrix to row echelon form U;
- the nonzero rows of U will form a basis for $\operatorname{row}(A)$;
- the number of nonzero rows of U equals $\operatorname{rank} A$.

Example 4.60. Let

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

Then

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

Thus

$$\left\{ \begin{pmatrix} 1 & -3 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 1 & -1 \end{pmatrix} \right\}$$

is a basis for $\operatorname{row}(A)$, and $\operatorname{rank} A = 2$.

It turns out that the rank of a matrix A is intimately connected with the dimension of its nullspace $N(A)$. Before formulating this relation, we require some more terminology:

Definition 4.61. If $A \in \mathbb{R}^{m\times n}$, then $\dim N(A)$ is called the **nullity** of A, and is denoted by $\operatorname{nul} A$.

Example 4.62. Find the nullity of the matrix

$$A = \begin{pmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{pmatrix}.$$

Solution. We have already calculated the nullspace $N(A)$ of this matrix in Example 4.16 by bringing A to row echelon form U and then using back substitution to solve $U\mathbf{x} = \mathbf{0}$, giving

$$N(A) = \{\, \alpha\mathbf{x}_1 + \beta\mathbf{x}_2 + \gamma\mathbf{x}_3 \mid \alpha, \beta, \gamma \in \mathbb{R} \,\} \,,$$

where

$$\mathbf{x}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \,, \quad \mathbf{x}_2 = \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} \,, \quad \mathbf{x}_3 = \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} \,.$$

It is not difficult to see that $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ are linearly independent, so $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is a basis for $N(A)$. Thus, $\operatorname{nul} A = 3$. □

Notice that in the above example the nullity of A is equal to the number of free variables of the system $A\mathbf{x} = \mathbf{0}$. This is no coincidence, but true in general! If this is not obvious to you, you should mull about this for a little while!

The connection between the rank and nullity of a matrix, alluded to above, is the content of the following beautiful theorem with an ugly name:

Theorem 4.63 (Rank-Nullity Theorem)**.** *If $A \in \mathbb{R}^{m\times n}$, then*

$$\operatorname{rank} A + \operatorname{nul} A = n \,.$$

Proof. Bring A to row echelon form U. Write $r = \operatorname{rank} A$. Now observe that U has r non-zero rows, hence $U\mathbf{x} = \mathbf{0}$ has $n - r$ free variables, so $\operatorname{nul} A = n - r$. □

We now return to the perhaps rather surprising connection between the dimensions of the row space and the column space of a matrix.

Theorem 4.64. *Let $A \in \mathbb{R}^{m\times n}$. Then*

$$\dim \operatorname{col}(A) = \dim \operatorname{row}(A) \,.$$

Proof. If A has rank r, then the row echelon form U of A will have r leading 1's. The columns of U corresponding to the leading 1's will be linearly independent. They do not, however, form a basis of $\operatorname{col}(A)$, since in general A and U will have different column spaces. Let U_L denote the matrix obtained from U by deleting all the columns corresponding to the free variables. Delete the same columns from A and denote the new matrix by A_L. The matrices A_L and U_L are row equivalent. Thus, if $\mathbf{x}$ is a solution to $A_L\mathbf{x} = \mathbf{0}$, then $U_L\mathbf{x} = \mathbf{0}$. Since the columns of U_L are linearly independent, $\mathbf{x}$ must equal $\mathbf{0}$. It follows that the columns of A_L must be linearly independent as well. Since A_L has r columns, the dimension of $\operatorname{row}(A)$ must be at least r, that is,

$$\dim \operatorname{col}(A) \geq r = \dim \operatorname{row}(A) \,.$$

Applying the above inequality to A^T yields

$$\dim \operatorname{row}(A) = \dim \operatorname{col}(A^T) \geq \dim \operatorname{row}(A^T) = \dim \operatorname{col}(A) \,,$$

so $\dim \operatorname{col}(A) = \dim \operatorname{row}(A)$ as claimed. □

The above proof shows how to find a basis for the column space of a matrix:

In order to find a basis for the column space of a matrix A:

- bring A to row echelon form and identify the leading variables;
- the columns of A containing the leading variables form a basis for $\mathrm{col}(A)$.

Example 4.65. Let

$$A = \begin{pmatrix} 1 & -1 & 3 & 2 & 1 \\ 1 & 0 & 1 & 4 & 1 \\ 2 & -1 & 4 & 7 & 4 \end{pmatrix}.$$

Then the row echelon form of A is

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

The leading variables are in columns 1,2, and 4. Thus a basis for $\mathrm{col}(A)$ is given by

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 7 \end{pmatrix} \right\}.$$

Chapter 5

Linear Transformations

5.1 Definition and examples

Linear transformations are the bread and butter of Linear Algebra. You have already encountered them in Geometry I. Roughly speaking a linear transformation is a mapping between two vector spaces that preserves the linear structure of the underlying spaces. To be precise:

Definition 5.1. Let V and W be two vector spaces. A mapping $L : V \to W$ (that is, a mapping from V to W) is said to be a **linear transformation** or a **linear mapping** if it satisfies the following two conditions:

(i) $L(\mathbf{v} + \mathbf{w}) = L(\mathbf{v}) + L(\mathbf{w})$ for all $\mathbf{v}$ and $\mathbf{w}$ in V;

(ii) $L(\alpha\mathbf{v}) = \alpha L(\mathbf{v})$ for all $\mathbf{v}$ in V and all scalars α.

Example 5.2. Let $L : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$L(\mathbf{x}) = 2\mathbf{x}\,.$$

Then L is linear since, if $\mathbf{x}$ and $\mathbf{y}$ are arbitrary vectors in $\mathbb{R}^2$ and α is an arbitrary real number, then

(i) $L(\mathbf{x} + \mathbf{y}) = 2(\mathbf{x} + \mathbf{y}) = 2\mathbf{x} + 2\mathbf{y} = L(\mathbf{x}) + L(\mathbf{y})$;

(ii) $L(\alpha\mathbf{x}) = 2(\alpha\mathbf{x}) = \alpha(2\mathbf{x}) = \alpha L(\mathbf{x})$.

Example 5.3. Let $L : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$L(\mathbf{x}) = x_1\mathbf{e}_1\,, \quad \text{where} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\,.$$

Then L is linear. In order to see this suppose that $\mathbf{x}$ and $\mathbf{y}$ are arbitrary vectors in $\mathbb{R}^2$ with

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\,, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\,.$$

Notice that, if α is an arbitrary real number, then

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} \quad \text{and} \quad \alpha\mathbf{x} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \end{pmatrix}\,.$$

Thus

(i) $L(\mathbf{x}+\mathbf{y}) = (x_1+y_1)\mathbf{e}_1 = x_1\mathbf{e}_1 + y_1\mathbf{e}_1 = L(\mathbf{x}) + L(\mathbf{y})$;

(ii) $L(\alpha\mathbf{x}) = (\alpha x_1)\mathbf{e}_1 = \alpha(x_1\mathbf{e}_1) = \alpha L(\mathbf{x})$.

Hence L is linear, as claimed.

In order to shorten statements of theorems and examples let us introduce the following convention:

If $\mathbf{x}$ is a vector in $\mathbb{R}^n$, we shall henceforth denote its i-th entry by x_i, and similarly for vectors in $\mathbb{R}^n$ denoted by other bold symbols. So, for example, if $\mathbf{y} = (1,4,2,7)^T \in \mathbb{R}^4$, then $y_3 = 2$.

Example 5.4. Let $L : \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$L(\mathbf{x}) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}.$$

L is linear, since, if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$, then

(i) $L(\mathbf{x}+\mathbf{y}) = \begin{pmatrix} -(x_2+y_2) \\ x_1+y_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} + \begin{pmatrix} -y_2 \\ y_1 \end{pmatrix} = L(\mathbf{x}) + L(\mathbf{y})$;

(ii) $L(\alpha\mathbf{x}) = \begin{pmatrix} -\alpha x_2 \\ \alpha x_1 \end{pmatrix} = \alpha \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} = \alpha L(\mathbf{x})$.

Example 5.5. The mapping $M : \mathbb{R}^2 \to \mathbb{R}^1$ defined by

$$M(\mathbf{x}) = \sqrt{x_1^2 + x_2^2}$$

is not linear. In order to see this note that $M((1,0)^T) = \sqrt{1^2} = 1$ while $M(-(1,0)^T) = M((-1,0)^T) = \sqrt{(-1)^2} = 1$. Thus

$$M(-(1,0)^T) = 1 \neq -1 = -M((1,0)^T).$$

Important Observation. Any $m \times n$ matrix A induces a linear transformation $L_A : \mathbb{R}^n \to \mathbb{R}^m$ given by

$$L_A(\mathbf{x}) = A\mathbf{x} \quad \text{for each } \mathbf{x} \in \mathbb{R}^n.$$

The transformation L_A is linear, since, if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, then

(i) $L_A(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} = L_A(\mathbf{x}) + L_A(\mathbf{y})$;

(ii) $L_A(\alpha\mathbf{x}) = A(\alpha\mathbf{x}) = \alpha A\mathbf{x} = \alpha L_A(\mathbf{x})$.

In other words, every $m \times n$ matrix gives rise to a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. We shall see shortly that, conversely, *every* linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ arises from an $m \times n$ matrix.

5.2 Linear transformations on general vector spaces

So far we have only considered linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$. In this short section, we shall have a look at linear transformations on abstract vector spaces. We start with some general observations concerning linear transformations on abstract vector spaces and finish with two examples.

Theorem 5.6. *If V and W are vector spaces and $L : V \to W$ is a linear transformation then:*

(a) $L(\mathbf{0}) = \mathbf{0}$*;*

(b) $L(-\mathbf{v}) = -L(\mathbf{v})$ *for any* $\mathbf{v} \in V$*;*

(c) $L(\sum_{i=1}^{n} \alpha_i \mathbf{v}_i) = \sum_{i=1}^{n} \alpha_i L(\mathbf{v}_i)$ *for any* $\mathbf{v}_i \in V$ *and any scalars* α_i *where* $i = 1, \ldots, n$*.*

Proof.

(a) $L(\mathbf{0}) = L(0\mathbf{0}) = 0L(\mathbf{0}) = \mathbf{0}$;

(b) $L(-\mathbf{v}) = L((-1)\mathbf{v}) = (-1)L(\mathbf{v}) = -L(\mathbf{v})$;

(c) follows by repeated application of the defining properties (i) and (ii) of linear transformations.

Note that we have used Theorem 4.5 for the proof of (a) and (b). □

Let's look at some examples, which should convince you that linear transformations arise naturally in other areas of Mathematics.

Example 5.7. Let $L : C[a, b] \to \mathbb{R}^1$ be defined by

$$L(\mathbf{f}) = \int_a^b \mathbf{f}(t)\, dt\,.$$

L is linear since, if $\mathbf{f}, \mathbf{g} \in C[a, b]$ and $\alpha \in \mathbb{R}$, then

(i) $L(\mathbf{f} + \mathbf{g}) = \displaystyle\int_a^b (\mathbf{f}(t) + \mathbf{g}(t))\, dt = \int_a^b \mathbf{f}(t)\, dt + \int_a^b \mathbf{g}(t)\, dt = L(\mathbf{f}) + L(\mathbf{g})$;

(ii) $L(\alpha\mathbf{f}) = \displaystyle\int_a^b (\alpha\mathbf{f}(t))\, dt = \alpha \int_a^b \mathbf{f}(t)\, dt = \alpha L(\mathbf{f})$.

In other words, integration is a linear transformation.

Example 5.8. Let $D : C^1[a, b] \to C[a, b]$ be defined to be the transformation that sends an $\mathbf{f} \in C^1[a, b]$ to its derivative $\mathbf{f}' \in C[a, b]$, that is,

$$D(\mathbf{f}) = \mathbf{f}'\,.$$

Then D is linear since, if $\mathbf{f}, \mathbf{g} \in C^1[a, b]$ and $\alpha \in \mathbb{R}$, then

(i) $D(\mathbf{f} + \mathbf{g}) = (\mathbf{f} + \mathbf{g})' = \mathbf{f}' + \mathbf{g}' = D(\mathbf{f}) + D(\mathbf{g})$;

(ii) $D(\alpha\mathbf{f}) = (\alpha\mathbf{f})' = \alpha\mathbf{f}' = \alpha D(\mathbf{f})$.

In other words, differentiation is a linear transformation.

Example 5.9. Let V be a vector space and let $Id : V \to V$ denote the **identity transformation** (or **identity** for short) on V, that is,

$$Id(\mathbf{v}) = \mathbf{v} \quad \text{for all } \mathbf{v} \in V.$$

The transformation Id is linear, since, if $\mathbf{v}, \mathbf{w} \in V$ and α is a scalar, then

(i) $Id(\mathbf{v} + \mathbf{w}) = \mathbf{v} + \mathbf{w} = Id(\mathbf{v}) + Id(\mathbf{w})$;

(ii) $Id(\alpha\mathbf{v}) = \alpha\mathbf{v} = \alpha Id(\mathbf{v})$.

5.3 Image and Kernel

We shall now discuss two useful notions, namely that of the 'image' and that of the 'kernel' of a linear transformation, that help to generalise two notions that you have already encountered in connection with matrices.

Definition 5.10. Let V and W be vector spaces, and let $L : V \to W$ be a linear transformation. The **kernel** of L, denoted by $\ker(L)$, is the subset of V given by

$$\ker(L) = \{\, \mathbf{v} \in V \mid L(\mathbf{v}) = \mathbf{0} \,\} .$$

Example 5.11. If $A \in \mathbb{R}^{m\times n}$ and L_A is the corresponding linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$, then

$$\ker(L_A) = N(A) ,$$

that is, the kernel of L_A is the nullspace of A.

The previous example shows that the kernel of a linear transformation is the natural generalisation of the nullspace of a matrix.

Definition 5.12. Let V and W be vector spaces. Let $L : V \to W$ be a linear transformation and let H be a subspace of V. The **image** of H (under L), denoted by $L(H)$, is the subset of W given by

$$L(H) = \{\, \mathbf{w} \in W \mid \mathbf{w} = L(\mathbf{v}) \text{ for some } \mathbf{v} \in H \,\} .$$

The image $L(V)$ of the entire vector space V under L is called the **range** of L.

Example 5.13. If $A \in \mathbb{R}^{m\times n}$ and L_A is the corresponding linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$, then

$$L_A(\mathbb{R}^n) = \mathrm{col}(A) ,$$

that is, the range of L_A is the column space of A.

The previous example shows that the range of a linear transformation is the natural generalisation of the column space of a matrix.

We saw previously that the nullspace and the column space of an $m \times n$ matrix are subspaces of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. The same is true for the abstract analogues introduced above.

Theorem 5.14. *Let V and W and be vector spaces. If $L : V \to W$ is a linear transformation and H is a subspace of V, then*

(a) $\ker(L)$ is a subspace of V;

(b) $L(H)$ is a subspace of W.

Proof.

(a) First observe that $\ker(L)$ is not empty since $\mathbf{0} \in \ker(L)$ by Theorem 5.6. Suppose now that $\mathbf{v}_1, \mathbf{v}_2 \in \ker(L)$. Then

$$L(\mathbf{v}_1 + \mathbf{v}_2) = L(\mathbf{v}_1) + L(\mathbf{v}_2) = \mathbf{0} + \mathbf{0} = \mathbf{0}\,,$$

so $\mathbf{v}_1 + \mathbf{v}_2 \in \ker(L)$. Moreover, if $\mathbf{v} \in \ker(L)$ and α is a scalar, then

$$L(\alpha\mathbf{v}) = \alpha L(\mathbf{v}) = \alpha\mathbf{0} = \mathbf{0}\,,$$

so $\alpha\mathbf{v} \in \ker(L)$. Thus, as $\ker(L)$ is closed under addition and scalar multiplication, it is a subspace of V as claimed.

(b) First observe that $L(H)$ is not empty since $\mathbf{0} \in L(H)$ by Theorem 5.6. Suppose now that $\mathbf{w}_1, \mathbf{w}_2 \in L(H)$. Then there are $\mathbf{v}_1, \mathbf{v}_2 \in H$ such that $L(\mathbf{v}_1) = \mathbf{w}_1$ and $L(\mathbf{v}_2) = \mathbf{w}_2$ and so

$$\mathbf{w}_1 + \mathbf{w}_2 = L(\mathbf{v}_1) + L(\mathbf{v}_2) = L(\mathbf{v}_1 + \mathbf{v}_2)\,.$$

But $\mathbf{v}_1 + \mathbf{v}_2 \in H$, because H is a subspace, so $\mathbf{w}_1 + \mathbf{w}_2 \in L(H)$. Moreover, if $\mathbf{w} \in L(H)$ and α is a scalar, then there is $\mathbf{v} \in H$ such that $L(\mathbf{v}) = \mathbf{w}$ and so

$$\alpha\mathbf{w} = \alpha L(\mathbf{v}) = L(\alpha\mathbf{v})\,.$$

But $\alpha\mathbf{v} \in H$, because H is a subspace, so $\alpha\mathbf{w} \in L(H)$. Thus, as $L(H)$ is closed under addition and scalar multiplication, it is a subspace of W as claimed.

□

Example 5.15. Let $D : P_3 \to P_3$ be the differentiation transformation given by

$$D(\mathbf{p}) = \mathbf{p}'\,.$$

Find $\ker(D)$ and $D(P_3)$.

Solution. The derivative of a polynomial $\mathbf{p} \in P_3$ is the zero polynomial if and only if $\mathbf{p}$ is a constant. Thus

$$\ker(D) = P_0\,.$$

Since differentiation lowers the degree of a polynomial by 1, we see that $D(P_3)$ is a subspace of P_2. However, any polynomial in P_2 has an antiderivative in P_3, so every polynomial in P_2 will be the image of a polynomial in P_3 under D. Thus

$$D(P_3) = P_2\,.$$

□

Example 5.16 (Optional; for those of you who have taken MTH4102 'Differential Equations')**.** Let $L : C^1[-1, 1] \to C[-1, 1]$ be the linear transformation given by

$$L(\mathbf{f}) = \mathbf{f} + \mathbf{f}'\,.$$

Find the kernel and range of L.

Solution. In order to determine the kernel of L we need to find all $\mathbf{f} \in C^1[-1,1]$ such that

$$\mathbf{f} + \mathbf{f}' = \mathbf{0}\,. \tag{5.1}$$

This is a first order homogeneous differential equation with integrating factor e^t. Thus, a function $\mathbf{f}$ satisfies (5.1) if and only if

$$\frac{d}{dt}(e^t\mathbf{f}(t)) = 0\,,$$

so

$$e^t\mathbf{f}(t) = \alpha\,,$$

for some $\alpha \in \mathbb{R}$, and hence

$$\mathbf{f}(t) = \alpha e^{-t}\,.$$

Thus

$$\ker(L) = \mathrm{Span}\,(\mathbf{h})\,,$$

where $\mathbf{h}(t) = e^{-t}$.

The moral of this part of the example is that whenever you are solving a linear homogeneous differential equation, you are in fact calculating the kernel of a certain linear transformation!

In order to determine the range of L, notice that L clearly sends continuously differentiable functions to continuous functions. The question is whether *every* continuous function arises as an image of some $\mathbf{f} \in C^1[-1,1]$ under L. The answer is yes! To see this, fix $\mathbf{g} \in C[-1,1]$. We need to show that there is an $\mathbf{f} \in C^1[-1,1]$ such that $L(\mathbf{f}) = \mathbf{g}$, that is,

$$\mathbf{f}' + \mathbf{f} = \mathbf{g}\,. \tag{5.2}$$

This is a first order linear inhomogeneous differential equation for $\mathbf{f}$. Using the integrating factor e^t we find that (5.2) is equivalent to

$$\frac{d}{dt}(e^t\mathbf{f}(t)) = e^t\mathbf{g}(t)\,.$$

But the right hand side of the equation above has an antiderivative, say $\mathbf{H}$, that is,

$$\frac{d}{dt}\mathbf{H}(t) = e^t\mathbf{g}(t)$$

so

$$e^t\mathbf{f}(t) = \mathbf{H}(t) + \alpha\,,$$

for some $\alpha \in \mathbb{R}$, hence

$$\mathbf{f}(t) = e^{-t}\mathbf{H}(t) + \alpha e^{-t}\,.$$

Notice that the $\mathbf{f}$ just found is clearly continuously differentiable. To summarise, we have shown that, given $\mathbf{g} \in C[-1,1]$, we can find $\mathbf{f} \in C^1[-1,1]$ such that $L(\mathbf{f}) = \mathbf{g}$. In other words, the range of L is $C[-1,1]$.

The moral of this part of the example is the following: the statement that the range of L is $C[-1,1]$ means that the differential equation (5.2) has a solution for any $\mathbf{g} \in C[-1,1]$. So statements about ranges of certain linear transformations are in fact statements about the existence of solutions for certain classes of differential equations!

□

5.4 Matrix representations of linear transformations

This subsection contains the central result of this chapter, if not of the whole module. In essence, this result says that every linear transformation between finite dimensional vector spaces can be viewed as a matrix. Or in other words, matrices provide concrete realisations of abstract linear transformations. This is one of the main reasons why matrices underpin large parts of contemporary Mathematics.

Let's start with a baby-version of this miraculous result. To appreciate it, recall the important observation that any matrix $A \in \mathbb{R}^{m\times n}$ gives rise to a linear transformation $L_A : \mathbb{R}^n \to \mathbb{R}^m$ by defining $L_A(\mathbf{x}) = A\mathbf{x}$. The following theorem says that *every* linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ arises in this way:

Theorem 5.17. *If L is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$, then there is an $m \times n$ matrix A such that*

$$L(\mathbf{x}) = A\mathbf{x} \quad \textit{for each } \mathbf{x} \in \mathbb{R}^n.$$

In fact, the j-th column $\mathbf{a}_j$ of A is given by

$$\mathbf{a}_j = L(\mathbf{e}_j) \quad \textit{for } j = 1, \dots, n\,,$$

where $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ denotes the standard basis of $\mathbb{R}^n$.

Proof. For $j = 1, \dots, n$ let $\mathbf{a}_j = L(\mathbf{e}_j)$ and let $A \in \mathbb{R}^{m\times n}$ be the matrix whose j-th column is $\mathbf{a}_j$. Now, if $\mathbf{x} \in \mathbb{R}^n$, then

$$\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n\,,$$

so

$$L(\mathbf{x}) = L(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) = x_1L(\mathbf{e}_1) + \cdots + x_nL(\mathbf{e}_n) = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = A\mathbf{x}\,.$$

□

To paraphrase the theorem above yet again: given a linear transformation L between $\mathbb{R}^n$ and $\mathbb{R}^m$ we can find a matrix that represents it. For this reason the matrix A constructed above is called a *matrix representation* of the linear transformation L. For the construction of A we have used the standard basis. We shall see shortly that there is nothing special about the standard basis, that is, we could have chosen any other basis to represent a given transformation. This will be the content of the 'Matrix Representation Theorem' we will encounter shortly.

Example 5.18. Let $L : \mathbb{R}^3 \to \mathbb{R}^2$ be given by

$$L(\mathbf{x}) = \begin{pmatrix} x_1 - x_2 \\ x_2 + 2x_3 \end{pmatrix}.$$

The transformation L is easily seen to be linear. Now

$$L(\mathbf{e}_1) = \begin{pmatrix} 1 & - & 0 \\ 0 & + & 2\cdot 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
$$L(\mathbf{e}_2) = \begin{pmatrix} 0 & - & 1 \\ 1 & + & 2\cdot 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$
$$L(\mathbf{e}_3) = \begin{pmatrix} 0 & - & 0 \\ 0 & + & 2\cdot 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

so if we set

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

then indeed

$$A\mathbf{x} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ x_2 + 2x_3 \end{pmatrix} = L(\mathbf{x}).$$

We are now ready for the main result alluded to at the beginning of this section. This truly marvellous result is an extension of Theorem 5.17. It states that any linear transformation between arbitrary finite dimensional vector spaces can be represented by a matrix.

In order to appreciate this statement, let's suppose that V is an n-dimensional vector space with basis $\mathcal{V}$ and that W is an m-dimensional vector space with basis $\mathcal{W}$. If $L : V \to W$ is a linear transformation we can ask the following question: given an arbitrary vector $\mathbf{v} \in V$ with $\mathcal{V}$-coordinates $[\mathbf{v}]_{\mathcal{V}}$, what are the $\mathcal{W}$-coordinates of $L(\mathbf{v})$? It turns out that there is an $m \times n$ matrix A, not depending on $\mathbf{v}$, such that

$$[L(\mathbf{v})]_{\mathcal{W}} = A[\mathbf{v}]_{\mathcal{V}}.$$

In other words, the matrix A represents the change of coordinates effected by the linear transformation L. This important fact is the content of the following theorem:

Theorem 5.19 (Matrix Representation Theorem)**.** *If $\mathcal{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $\mathcal{W} = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ are bases for vector spaces V and W respectively, then corresponding to each linear transformation $L : V \to W$ there is an $m \times n$ matrix A such that*

$$[L(\mathbf{v})]_{\mathcal{W}} = A[\mathbf{v}]_{\mathcal{V}} \quad \text{for each } \mathbf{v} \in V.$$

In fact, the j-th column $\mathbf{a}_j$ of A is given by

$$\mathbf{a}_j = [L(\mathbf{v}_j)]_{\mathcal{W}}.$$

Proof. For each $j = 1, \ldots n$ let $\mathbf{a}_j = (a_{1j}, \ldots, a_{mj})^T$ be the $\mathcal{W}$-coordinate vector of $L(\mathbf{v}_j)$. Thus

$$L(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + \cdots + a_{mj}\mathbf{w}_m = \sum_{i=1}^{m} a_{ij}\mathbf{w}_i.$$

Fix $\mathbf{v} \in V$. Let $\mathbf{x}$ be the $\mathcal{V}$-coordinate vector of $\mathbf{v}$, that is,

$$\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n = \sum_{j=1}^{n} x_j\mathbf{v}_j.$$

Then

$$\begin{aligned} L(\mathbf{v}) &= L(\sum_{j=1}^{n} x_j\mathbf{v}_j) \\ &= \sum_{j=1}^{n} x_j L(\mathbf{v}_j) \\ &= \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{m} a_{ij}\mathbf{w}_i \right) \\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}x_j \right) \mathbf{w}_i. \end{aligned}$$

If we now call

$$y_i = \sum_{j=1}^{n} a_{ij}x_j \quad \text{for } i = 1, \ldots, m\,,$$

then the above calculation shows that

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = A\mathbf{x}$$

is the $\mathcal{W}$-coordinate vector of $L(\mathbf{v})$, that is

$$[L(\mathbf{v})]_{\mathcal{W}} = A[\mathbf{v}]_{\mathcal{V}}\,.$$

□

Definition 5.20. Given vector spaces V and W with corresponding bases $\mathcal{V}$ and $\mathcal{W}$, and a linear transformation $L : V \to W$, we call the matrix A constructed in the theorem above the **matrix representation of L with respect to $\mathcal{V}$ and $\mathcal{W}$**, and denote it by $\left[L\right]_{\mathcal{W}}^{\mathcal{V}}$. Thus, for any $\mathbf{v} \in V$ we have

$$[L(\mathbf{v})]_{\mathcal{W}} = \left[L\right]_{\mathcal{W}}^{\mathcal{V}}[\mathbf{v}]_{\mathcal{V}}\,.$$

Example 5.21. Let $L : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by

$$L(\mathbf{x}) = x_1\mathbf{b}_1 + (x_2 + x_3)\mathbf{b}_2\,,$$

where

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Find the matrix representation of L with respect to the standard basis $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and the basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

Solution. Since

$$\begin{array}{rcrcr} L(\mathbf{e}_1) & = & 1\mathbf{b}_1 & + & 0\mathbf{b}_2 \\ L(\mathbf{e}_2) & = & 0\mathbf{b}_1 & + & 1\mathbf{b}_2 \\ L(\mathbf{e}_3) & = & 0\mathbf{b}_1 & + & 1\mathbf{b}_2 \end{array}$$

we see that

$$[L(\mathbf{e}_1)]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad [L(\mathbf{e}_2)]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad [L(\mathbf{e}_3)]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

so

$$\left[L\right]_{\mathcal{B}}^{\mathcal{E}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

□

Example 5.22. Consider the linear transformation $D : P_2 \to P_1$ given by

$$(D(\mathbf{p}))(t) = \mathbf{p}'(t)\,.$$

Define

$$\mathbf{p}_1(t) = 1, \quad \mathbf{p}_2(t) = t, \quad \mathbf{p}_3(t) = t^2,$$

and let $\mathcal{P}_2 = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ and $\mathcal{P}_1 = \{\mathbf{p}_1, \mathbf{p}_2\}$ be bases for P_2 and P_1 respectively. Since

$$\begin{aligned}(D(\mathbf{p}_1))(t) &= \mathbf{p}_1'(t) = 0\\ (D(\mathbf{p}_2))(t) &= \mathbf{p}_2'(t) = 1\\ (D(\mathbf{p}_3))(t) &= \mathbf{p}_3'(t) = 2t\end{aligned}$$

we have

$$\begin{aligned}D(\mathbf{p}_1) &= 0\mathbf{p}_1 + 0\mathbf{p}_2\\ D(\mathbf{p}_2) &= 1\mathbf{p}_1 + 0\mathbf{p}_2\\ D(\mathbf{p}_3) &= 0\mathbf{p}_1 + 2\mathbf{p}_2\end{aligned}$$

so

$$\big[D\big]_{\mathcal{P}_1}^{\mathcal{P}_2} = \begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 2\end{pmatrix}.$$

Suppose now that $\mathbf{p} \in P_2$ is given by

$$\mathbf{p}(t) = a + bt + ct^2\,.$$

We want to find $D(\mathbf{p})$. Of course we could do this working directly from the definition of D, but we can also use the Matrix Representation Theorem: since

$$\mathbf{p} = a\mathbf{p}_1 + b\mathbf{p}_2 + c\mathbf{p}_3\,,$$

we have, by the Matrix Representation Theorem,

$$[D(\mathbf{p})]_{\mathcal{P}_1} = \big[D\big]_{\mathcal{P}_1}^{\mathcal{P}_2}[\mathbf{p}]_{\mathcal{P}_2} = \begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 2\end{pmatrix}\begin{pmatrix}a\\ b\\ c\end{pmatrix} = \begin{pmatrix}b\\ 2c\end{pmatrix},$$

so

$$D(\mathbf{p}) = b\mathbf{p}_1 + 2c\mathbf{p}_2\,,$$

that is,

$$\mathbf{p}'(t) = b + 2ct\,,$$

as expected.

Example 5.23. Let $A \in \mathbb{R}^{m\times n}$, and let L_A be the corresponding linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. Since $L_A(\mathbf{e}_j)$ is just the j-th column of A, we see that the matrix representation of L_A with respect to the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$ is just A itself.

5.5 Composition of linear transformations

In this short section we shall have look at composition of linear transformations and its effect on the corresponding matrix representations. Suppose that U, V and W are vector spaces and that we are given two linear transformations

$$\begin{aligned}T &: U \to V\,,\\ S &: V \to W\,.\end{aligned}$$

We can then form a new transformation $S \circ T : U \to W$ by defining

$$(S \circ T)(\mathbf{u}) = S(T(\mathbf{u})) \quad \text{for each } \mathbf{u} \in U.$$

The transformation $S \circ T$ is called the **composite** of S and T. Observe that $S \circ T$ is linear as well. In order to see this, let $\mathbf{u}_1, \mathbf{u}_2 \in U$ and α_1, α_2 be scalars. Then

$$\begin{aligned}(S \circ T)(\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2) &= S(T(\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2)) \\ &= S(\alpha_1 T(\mathbf{u}_1) + \alpha_2 T(\mathbf{u}_2)) \\ &= \alpha_1 S(T(\mathbf{u}_1)) + \alpha_2 S(T(\mathbf{u}_2)) \\ &= \alpha_1 (S \circ T)(\mathbf{u}_1) + \alpha_2 (S \circ T)(\mathbf{u}_2)\,.\end{aligned}$$

Choosing $\alpha_1 = \alpha_2 = 1$ in the above equality gives

$$(S \circ T)(\mathbf{u}_1 + \mathbf{u}_2) = (S \circ T)(\mathbf{u}_1) + (S \circ T)(\mathbf{u}_2)\,,$$

while choosing $\alpha_1 = 1$ and $\alpha_2 = 0$ gives

$$(S \circ T)(\alpha_1 \mathbf{u}_1) = \alpha_1 (S \circ T)(\mathbf{u}_1)\,,$$

so $S \circ T$ is linear, as claimed.

Example 5.24. Suppose that $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times r}$. Let $L_A : \mathbb{R}^n \to \mathbb{R}^m$ and $L_B : \mathbb{R}^r \to \mathbb{R}^n$ be the corresponding linear transformations. Then $L_A \circ L_B : \mathbb{R}^r \to \mathbb{R}^m$ is the linear transformation given by

$$(L_A \circ L_B)(\mathbf{x}) = L_A(L_B(\mathbf{x})) = L_A(B\mathbf{x}) = AB\mathbf{x}\,,$$

so

$$L_A \circ L_B = L_{AB}\,.$$

In other words, the composite of L_A and L_B is the linear transformation arising from the product AB. Ultimately, this is the deeper reason why matrix multiplication is defined as it is.

We shall shortly discuss a more abstract analogue of the above example for two arbitrary linear transformations between finite dimensional vector spaces. Before doing so we digress to introduce some more notation:

Notation 5.25. If V is a vector space and $L : V \to V$ is linear, we can compose L with itself and call the resulting linear transformation L^2, that is,

$$L^2 = L \circ L\,.$$

More generally, we write

$$L^n = \underbrace{L \circ \cdots \circ L}_{n \text{ compositions}}$$

for the n-fold composite of L with itself.

Example 5.26. Let $D : P_5 \to P_5$ be the linear transformation given by

$$D(\mathbf{p}) = \mathbf{p}'\,.$$

Then

$$\begin{aligned}D^2(\mathbf{p}) &= \mathbf{p}'' \\ D^3(\mathbf{p}) &= \mathbf{p}'''\end{aligned}$$

and so forth.

Suppose now that U, V and W are finite dimensional vector spaces with corresponding bases $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$. If we are given two linear transformations $T : U \to V$ and $S : V \to W$ we may want to ask the question, how the matrix representations of S and T are related to that of $S \circ T$. The following theorem provides a neat answer:

Theorem 5.27 (Composition Formula)**.** *Let U, V and W be finite dimensional vector spaces with corresponding bases $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$. Suppose that $T : U \to V$ and $S : V \to W$ are linear transformations. Then*

$$\big[S \circ T\big]_{\mathcal{W}}^{\mathcal{U}} = \big[S\big]_{\mathcal{W}}^{\mathcal{V}} \big[T\big]_{\mathcal{V}}^{\mathcal{U}} ,$$

that is, the matrix representation of $S \circ T$ with respect to $\mathcal{U}$ and $\mathcal{W}$ is the matrix product of the matrix representation of S with respect $\mathcal{V}$ and $\mathcal{W}$ and the matrix representation of T with respect to $\mathcal{U}$ and $\mathcal{V}$.

Proof. Write $A = \big[S\big]_{\mathcal{W}}^{\mathcal{V}}$ and $B = \big[T\big]_{\mathcal{V}}^{\mathcal{U}}$. Let $\mathbf{u} \in U$ and let $\mathbf{v} = T(\mathbf{u}) \in V$. Thus, by the Matrix Representation Theorem,

$$[S(\mathbf{v})]_{\mathcal{W}} = A[\mathbf{v}]_{\mathcal{V}} ,$$
$$[T(\mathbf{u})]_{\mathcal{V}} = B[\mathbf{u}]_{\mathcal{U}} .$$

Hence

$$[(S \circ T)(\mathbf{u})]_{\mathcal{W}} = [S(T(\mathbf{u}))]_{\mathcal{W}} = [S(\mathbf{v})]_{\mathcal{W}} = A[\mathbf{v}]_{\mathcal{V}} = A[T(\mathbf{u})]_{\mathcal{V}} = AB[\mathbf{u}]_{\mathcal{U}} .$$

Since $\mathbf{u} \in U$ was arbitrary, the above equation means that AB must be the matrix representation of $S \circ T$ with respect to $\mathcal{U}$ and $\mathcal{W}$, that is,

$$AB = \big[S \circ T\big]_{\mathcal{W}}^{\mathcal{U}} .$$

□

In essence, the Composition Formula states that composition of linear transformations corresponds to matrix multiplication of the corresponding matrix representations. We shall see some applications of this formula in the next section. For the moment we shall be content with the following simple consequence:

Corollary 5.28. *Let V be a finite dimensional vector space with basis $\mathcal{V}$ and let $L : V \to V$. If A is the matrix representation of L with respect to $\mathcal{V}$, then A^n is the matrix representation of L^n with respect to $\mathcal{V}$, that is*

$$\big[L^n\big]_{\mathcal{V}}^{\mathcal{V}} = A^n .$$

5.6 Change of basis and similarity

In this last section of the chapter we shall consider the problem of how the matrix representation of a given linear transformation changes when the bases of the underlying vector spaces are changed. The solution to this problem will come in handy in the very last chapter of these notes.

Before turning our attention to this problem we shall first consider the slightly simpler one of how to describe coordinate changes in abstract vector spaces, thus extending the discussion of coordinate changes in $\mathbb{R}^n$ detailed in Chapter 4.

Question. Let V be a finite dimensional vector space with bases $\mathcal{U}$ and $\mathcal{W}$.

Given $\mathbf{v} \in V$ and its $\mathcal{U}$-coordinate vector $[\mathbf{v}]_{\mathcal{U}}$ how do we find the $\mathcal{W}$-coordinate vector of $\mathbf{v}$?

The answer is the content of the following theorem.

Theorem 5.29. *Let V be an n-dimensional vector space with bases $\mathcal{U} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ and $\mathcal{W} = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$. Then there is an $n \times n$ matrix S such that*

$$[\mathbf{v}]_{\mathcal{W}} = S[\mathbf{v}]_{\mathcal{U}} \quad \text{for any } \mathbf{v} \in V\,.$$

In fact, S is the matrix representation of the identity $Id : V \to V$ with respect to $\mathcal{U}$ and $\mathcal{W}$. In particular, the j-th column of S is given by $[\mathbf{u}_j]_{\mathcal{W}}$, where $j = 1, \ldots, n$.

Proof. Let S be the matrix representation of the identity $Id : V \to V$ with respect to $\mathcal{U}$ and $\mathcal{W}$. Then, by the Matrix Representation Theorem

$$[\mathbf{v}]_{\mathcal{W}} = [Id(\mathbf{v})]_{\mathcal{W}} = \left[Id\right]_{\mathcal{W}}^{\mathcal{U}} [\mathbf{v}]_{\mathcal{U}} = S[\mathbf{v}]_{\mathcal{W}}\,.$$

□

The $n \times n$ matrix S constructed in the theorem above is called the **transition matrix from $\mathcal{U}$ to $\mathcal{W}$**. It allows you to calculate the $\mathcal{W}$-coordinates of a vector, given that you know its $\mathcal{U}$-coordinates. It shouldn't come as a surprise that a transition matrix is invertible:

Theorem 5.30. *Let V be a finite dimensional vector space with bases $\mathcal{U}$ and $\mathcal{W}$. Then the transition matrix from $\mathcal{U}$ to $\mathcal{W}$ is invertible, and its inverse is the transition matrix from $\mathcal{W}$ to $\mathcal{U}$.*

Proof. By the Composition Formula we have

$$\begin{aligned} \left[Id\right]_{\mathcal{W}}^{\mathcal{U}} \left[Id\right]_{\mathcal{U}}^{\mathcal{W}} &= \left[Id \circ Id\right]_{\mathcal{W}}^{\mathcal{W}} = \left[Id\right]_{\mathcal{W}}^{\mathcal{W}} = I\,, \\ \left[Id\right]_{\mathcal{U}}^{\mathcal{W}} \left[Id\right]_{\mathcal{W}}^{\mathcal{U}} &= \left[Id \circ Id\right]_{\mathcal{U}}^{\mathcal{U}} = \left[Id\right]_{\mathcal{U}}^{\mathcal{U}} = I\,, \end{aligned}$$

so $\left[Id\right]_{\mathcal{W}}^{\mathcal{U}}$ is invertible with inverse $\left[Id\right]_{\mathcal{U}}^{\mathcal{W}}$. □

Example 5.31. Let $\mathcal{P} = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ and $\mathcal{Q} = \{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$ be the bases of P_2 given by

$$\begin{array}{lll} \mathbf{p}_1(t) = 1, & \mathbf{p}_2(t) = t, & \mathbf{p}_3(t) = t^2; \\ \mathbf{q}_1(t) = 1 + t^2, & \mathbf{q}_2(t) = t, & \mathbf{q}_3(t) = 2 + 3t^2. \end{array}$$

Now

$$\begin{array}{rcccccc} \mathbf{q}_1 &=& 1\mathbf{p}_1 &+& 0\mathbf{p}_2 &+& 1\mathbf{p}_3 \\ \mathbf{q}_2 &=& 0\mathbf{p}_1 &+& 1\mathbf{p}_2 &+& 0\mathbf{p}_3 \\ \mathbf{q}_3 &=& 2\mathbf{p}_1 &+& 0\mathbf{p}_2 &+& 3\mathbf{p}_3 \end{array}$$

so

$$[\mathbf{q}_1]_{\mathcal{P}} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad [\mathbf{q}_2]_{\mathcal{P}} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad [\mathbf{q}_3]_{\mathcal{P}} = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix},$$

hence the transition matrix from $\mathcal{Q}$ to $\mathcal{P}$ is

$$\left[Id\right]_{\mathcal{P}}^{\mathcal{Q}} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{pmatrix}.$$

Thus, the transition matrix from $\mathcal{P}$ to $\mathcal{Q}$ is

$$\big[Id\big]_{\mathcal{Q}}^{\mathcal{P}} = \Big(\big[Id\big]_{\mathcal{P}}^{\mathcal{Q}}\Big)^{-1} = \begin{pmatrix} 3 & 0 & -2 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} .$$

Now if $\mathbf{p}$ is given by

$$\mathbf{p}(t) = 2 - 3t + t^2 \,,$$

then $[\mathbf{p}]_{\mathcal{P}} = (2, -3, 1)^T$, so

$$[\mathbf{p}]_{\mathcal{Q}} = \big[Id\big]_{\mathcal{Q}}^{\mathcal{P}} [\mathbf{p}]_{\mathcal{P}} = \begin{pmatrix} 3 & 0 & -2 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ -3 \\ -1 \end{pmatrix} .$$

Thus

$$\mathbf{p} = 4\mathbf{q}_1 - 3\mathbf{q}_2 - \mathbf{q}_3 \,.$$

We now turn to the problem mentioned at the beginning of this section.

Question. Let V be a finite dimensional vector space with bases $\mathcal{U}$ and $\mathcal{W}$. Suppose that $L : V \to V$ is a linear transformation.
How are the two matrix representations $\big[L\big]_{\mathcal{U}}^{\mathcal{U}}$ and $\big[L\big]_{\mathcal{W}}^{\mathcal{W}}$ related?

The answer is contained in the following theorem:

Theorem 5.32. *Let V be a finite dimensional vector space with bases $\mathcal{U}$ and $\mathcal{W}$, and let $L : V \to V$ be a linear transformation. If S is the transition matrix from $\mathcal{U}$ to $\mathcal{W}$, then*

$$\big[L\big]_{\mathcal{U}}^{\mathcal{U}} = S^{-1}\big[L\big]_{\mathcal{W}}^{\mathcal{W}} S \,.$$

Proof. Note that $S = \big[Id\big]_{\mathcal{W}}^{\mathcal{U}}$ and $S^{-1} = \big[Id\big]_{\mathcal{U}}^{\mathcal{W}}$. Thus, by the composition formula,

$$S^{-1}\big[L\big]_{\mathcal{W}}^{\mathcal{W}} S = \big[Id\big]_{\mathcal{U}}^{\mathcal{W}} \big[L\big]_{\mathcal{W}}^{\mathcal{W}} \big[Id\big]_{\mathcal{W}}^{\mathcal{U}} = \big[Id\big]_{\mathcal{U}}^{\mathcal{W}} \big[L \circ Id\big]_{\mathcal{W}}^{\mathcal{U}} = \big[Id \circ L \circ Id\big]_{\mathcal{U}}^{\mathcal{U}} = \big[L\big]_{\mathcal{U}}^{\mathcal{U}} .$$

□

Example 5.33. Let $L : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation defined by $L(\mathbf{x}) = A\mathbf{x}$, where

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

(a) Find the matrix representation of L with respect to the basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$, where

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{b}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} .$$

(b) Find $L^n(\mathbf{x})$, where $\mathbf{x} = (-1, 2, 2)^T$.

Solution.

(a) First observe that the matrix representation of L with respect to the standard basis $\mathcal{E}$ of $\mathbb{R}^3$ is A itself. Now, the transition matrix S from $\mathcal{B}$ to $\mathcal{E}$ is simply

$$S = \big[Id\big]_{\mathcal{E}}^{\mathcal{B}} = P_{\mathcal{B}} = \begin{pmatrix} 1 & -2 & 1 \\ -1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} ,$$

so the transition matrix from $\mathcal{E}$ to $\mathcal{B}$ is

$$S^{-1} = \frac{1}{3} \begin{pmatrix} 0 & -3 & 3 \\ -1 & -1 & 2 \\ 1 & 1 & 1 \end{pmatrix} .$$

Thus

$$\big[L\big]_{\mathcal{B}}^{\mathcal{B}} = S^{-1}AS = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix} .$$

(b) Notice that it would be extremely silly to try and work out A^n in order to obtain $L^n(\mathbf{x}) = A^n\mathbf{x}$. What you should observe instead is that the matrix representation of L with respect to $\mathcal{B}$, call it B, is extremely simple: it is diagonal. Thus

$$\big[L^n\big]_{\mathcal{B}}^{\mathcal{B}} = B^n = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4^n \end{pmatrix} .$$

Hence, in order to work out $L^n(\mathbf{x})$ we only need to find the coordinates of $\mathbf{x}$ with respect to $\mathcal{B}$ and we are done. Now

$$[\mathbf{x}]_{\mathcal{B}} = \big[Id\big]_{\mathcal{B}}^{\mathcal{E}} [\mathbf{x}]_{\mathcal{E}} = S^{-1}\mathbf{x} = \frac{1}{3} \begin{pmatrix} 0 & -3 & 3 \\ -1 & -1 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} ,$$

so

$$[L^n(\mathbf{x})]_{\mathcal{B}} = \big[L^n\big]_{\mathcal{B}}^{\mathcal{B}} [\mathbf{x}]_{\mathcal{B}} = B^n \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 4^n \end{pmatrix} ,$$

and thus

$$L^n(\mathbf{x}) = \mathbf{b}_2 + 4^n\mathbf{b}_3 = \begin{pmatrix} -2 + 4^n \\ 1 + 4^n \\ 1 + 4^n \end{pmatrix} .$$

□

Let's have another look at the previous theorem: is states that if V is an n-dimensional vector space with two bases $\mathcal{U}$ and $\mathcal{W}$ and L is a linear transformation from V to itself, then the matrix representation of L with respect to $\mathcal{W}$, call it A, and the matrix representation of L with respect to $\mathcal{U}$, call it B, satisfy

$$B = S^{-1}AS\,,$$

where S is an invertible $n \times n$ matrix. This type of relation is of sufficient importance to merit a name of its own:

Definition 5.34. Let A and B be two $n \times n$ matrices. The matrix B is said to be **similar to** A if there is an invertible $S \in \mathbb{R}^{n \times n}$ such that

$$B = S^{-1}AS\,.$$

Notice that if B is similar to A, then A is similar to B, because if $R = S^{-1}$, then

$$A = SBS^{-1} = R^{-1}BR\,.$$

Thus we may simply say that A and B are similar matrices.

The content of Theorem 5.32 can now be rephrased as follows: if A and B are two matrix representations of the same linear transformation on a vector space, then A and B are similar.

Chapter 6

Orthogonality

In this chapter we will return to the concrete vector space $\mathbb{R}^n$ and add a new concept that will reveal new aspects of it. The added spice in the discussion is the notion of 'orthogonality'. This concept extends our intuitive notion of perpendicularity in $\mathbb{R}^2$ and $\mathbb{R}^3$ to $\mathbb{R}^n$. Innocent as it may seem, this new concept turns out to be a rather powerful device, as we shall see shortly.

6.1 Definition

We start by revisiting a concept that you have already encountered in Geometry I. Before stating it recall that a vector $\mathbf{x}$ in $\mathbb{R}^n$ is, by definition, an $n \times 1$ matrix. Given another vector $\mathbf{y}$ in $\mathbb{R}^n$, we may then form the matrix product $\mathbf{x}^T\mathbf{y}$ of the $1 \times n$ matrix $\mathbf{x}^T$ and the $n \times 1$ matrix $\mathbf{y}$. Notice that by the rules of matrix multiplication $\mathbf{x}^T\mathbf{y}$ is a 1×1 matrix, which we can simply think of as a real number.

Definition 6.1. Let $\mathbf{x}$ and $\mathbf{y}$ be two vectors in $\mathbb{R}^n$. The scalar $\mathbf{x}^T\mathbf{y}$ is called the **scalar product** or **dot product** of $\mathbf{x}$ and $\mathbf{y}$, and is often written $\mathbf{x}\cdot\mathbf{y}$. Thus, if

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

then

$$\mathbf{x}\cdot\mathbf{y} = \mathbf{x}^T\mathbf{y} = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = x_1y_1 + x_2y_2 + \cdots + x_ny_n\,.$$

Example 6.2. If

$$\mathbf{x} = \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix},$$

then

$$\mathbf{x}\cdot\mathbf{y} = \mathbf{x}^T\mathbf{y} = \begin{pmatrix}2 & -3 & 1\end{pmatrix}\begin{pmatrix}4\\5\\6\end{pmatrix} = 2\cdot 4 + (-3)\cdot 5 + 1\cdot 6 = 8 - 15 + 6 = -1\,,$$

$$\mathbf{y}\cdot\mathbf{x} = \mathbf{y}^T\mathbf{x} = \begin{pmatrix}4 & 5 & 6\end{pmatrix}\begin{pmatrix}2\\-3\\1\end{pmatrix} = 4\cdot 2 + 5\cdot(-3) + 6\cdot 1 = 8 - 15 + 6 = -1\,.$$

Having had a second look at the example above it should be clear why $\mathbf{x}\cdot\mathbf{y} = \mathbf{y}\cdot\mathbf{x}$. In fact, this is true in general. The following further properties of the dot product follow easily from properties of the transpose operation:

Theorem 6.3. *Let $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ be vectors in $\mathbb{R}^n$, and let α be a scalar. Then*

(a) $\mathbf{x}\cdot\mathbf{y} = \mathbf{y}\cdot\mathbf{x}$*;*

(b) $(\mathbf{x}+\mathbf{y})\cdot\mathbf{z} = \mathbf{x}\cdot\mathbf{z} + \mathbf{y}\cdot\mathbf{z}$*;*

(c) $(\alpha\mathbf{x})\cdot\mathbf{y} = \alpha(\mathbf{x}\cdot\mathbf{y}) = \mathbf{x}\cdot(\alpha\mathbf{y})$*;*

(d) $\mathbf{x}\cdot\mathbf{x} \geq 0$*, and* $\mathbf{x}\cdot\mathbf{x} = 0$ *if and only if* $\mathbf{x} = \mathbf{0}$*.*

Definition 6.4. If $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$, the **length** or **norm** of $\mathbf{x}$ is the nonnegative scalar $\|\mathbf{x}\|$ defined by

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}\cdot\mathbf{x}} = \sqrt{x_1^2 + \cdots + x_n^2}\,.$$

A vector whose length is 1 is called a **unit vector**.

Example 6.5. If $\mathbf{x} = (a, b)^T \in \mathbb{R}^2$, then

$$\|\mathbf{x}\| = \sqrt{a^2 + b^2}\,.$$

The above example should convince you that in $\mathbb{R}^2$ and $\mathbb{R}^3$ the definition of the length of a vector $\mathbf{x}$ coincides with the standard notion of the length of the line segment from the origin to $\mathbf{x}$.

Note that if $\mathbf{x} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ then

$$\|\alpha\mathbf{x}\| = |\alpha|\,\|\mathbf{x}\|\,,$$

because $\|\alpha\mathbf{x}\|^2 = (\alpha\mathbf{x})\cdot(\alpha\mathbf{x}) = \alpha^2(\mathbf{x}\cdot\mathbf{x}) = \alpha^2\|\mathbf{x}\|^2$. Thus, if $\mathbf{x} \neq \mathbf{0}$, we can always find a unit vector $\mathbf{y}$ in the same direction as $\mathbf{x}$ by setting

$$\mathbf{y} = \frac{1}{\|\mathbf{x}\|}\mathbf{x}\,.$$

The process of creating a unit vector $\mathbf{y}$ from $\mathbf{x}$ is called **normalising** $\mathbf{x}$.

Definition 6.6. For $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, the **distance between $\mathbf{x}$ and $\mathbf{y}$**, written $\mathrm{dist}(\mathbf{x}, \mathbf{y})$, is the length of $\mathbf{x} - \mathbf{y}$, that is,

$$\mathrm{dist}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|\,.$$

Again, you might want to mull for a second about this definition and convince yourself that in $\mathbb{R}^2$ and $\mathbb{R}^3$ the above definition of the distance between two vectors $\mathbf{x}$ and $\mathbf{y}$ coincides with the standard notion of the Euclidean distance between the two points represented by $\mathbf{x}$ and $\mathbf{y}$.

We now come to the main part of this introductory section. In essence, we will extend the concept of perpendicularity familiar in $\mathbb{R}^2$ and $\mathbb{R}^3$ to vectors in $\mathbb{R}^n$. In order to motivate it consider two lines through the origin in $\mathbb{R}^2$ or $\mathbb{R}^3$ determined by two vectors $\mathbf{x}$ and $\mathbf{y}$. By drawing a picture (which I will try to supply before long) of the triangle formed by the points representing $\mathbf{x}$, $\mathbf{y}$ and $-\mathbf{y}$, convince yourself that the lines in the direction of $\mathbf{x}$ and $\mathbf{y}$ are geometrically perpendicular if and only if the distance from $\mathbf{x}$ to $\mathbf{y}$ equals the distance from $\mathbf{x}$ to $-\mathbf{y}$. Now, using Theorem 6.3 repeatedly,

$$\begin{aligned}\operatorname{dist}(\mathbf{x},-\mathbf{y})^2 &= \|\mathbf{x}+\mathbf{y}\|^2\\ &= (\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}+\mathbf{y})\\ &= \mathbf{x}\cdot(\mathbf{x}+\mathbf{y})+\mathbf{y}\cdot(\mathbf{x}+\mathbf{y})\\ &= \mathbf{x}\cdot\mathbf{x}+\mathbf{x}\cdot\mathbf{y}+\mathbf{y}\cdot\mathbf{x}+\mathbf{y}\cdot\mathbf{y}\\ &= \|\mathbf{x}\|^2+\|\mathbf{y}\|^2+2\mathbf{x}\cdot\mathbf{y}\,.\end{aligned}$$

Using the same calculation with $-\mathbf{y}$ in place of $\mathbf{y}$ we see that

$$\begin{aligned}\operatorname{dist}(\mathbf{x},\mathbf{y})^2 &= \|\mathbf{x}\|^2+\|-\mathbf{y}\|^2+2\mathbf{x}\cdot(-\mathbf{y})\\ &= \|\mathbf{x}\|^2+\|\mathbf{y}\|^2-2\mathbf{x}\cdot\mathbf{y}\,.\end{aligned}$$

Thus

$$\operatorname{dist}(\mathbf{x},-\mathbf{y})=\operatorname{dist}(\mathbf{x},\mathbf{y}) \quad \text{if and only if} \quad \mathbf{x}\cdot\mathbf{y}=0\,.$$

In other words, the lines in the direction of $\mathbf{x}$ and $\mathbf{y}$ are geometrically perpendicular if and only if $\mathbf{x}\cdot\mathbf{y}=0$. We now take this observation as our starting point for the following definition, which extends the concept of perpendicularity to $\mathbb{R}^n$.

Definition 6.7. Two vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ are **orthogonal** (to each other) if $\mathbf{x}\cdot\mathbf{y}=0$.

Note that the zero vector is orthogonal to every other vector in $\mathbb{R}^n$.

The following theorem is an old acquaintance in new clothing, and, at the same time, contains a key fact about orthogonal vectors:

Theorem 6.8 (Pythagorean Theorem)**.** *Two vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ are orthogonal if and only if*

$$\|\mathbf{x}+\mathbf{y}\|^2=\|\mathbf{x}\|^2+\|\mathbf{y}\|^2\,.$$

Proof. See Exercise 3 on Coursework 9. □

6.2 Orthogonal complements

In this short section we introduce an important concept that will form the basis of subsequent developments.

Definition 6.9. Let Y be a subspace of $\mathbb{R}^n$. A vector $\mathbf{x}\in\mathbb{R}^n$ is said to be **orthogonal to** Y if $\mathbf{x}$ is orthogonal to every vector in Y. The set of all vectors in $\mathbb{R}^n$ that are orthogonal to Y is called the **orthogonal complement of** Y and is denoted by $Y^\perp$ (pronounced 'Y perpendicular' or 'Y perp' for short). Thus

$$Y^\perp=\{\,\mathbf{x}\in\mathbb{R}^n \mid \mathbf{x}\cdot\mathbf{y}=0 \text{ for all } \mathbf{y}\in Y\,\}\,.$$

Example 6.10. Let W be a plane through the origin in $\mathbb{R}^3$ and let L be the line through the origin and perpendicular to W. By construction, each vector in W is orthogonal to every vector in L, and each vector in L is orthogonal to every vector in W. Hence

$$L^\perp = W \quad \text{and} \quad W^\perp = L\,.$$

The following theorem collects some useful facts about orthogonal complements.

Theorem 6.11. *Let Y be a subspace of $\mathbb{R}^n$. Then:*

(a) $Y^\perp$ is a subspace of $\mathbb{R}^n$.

(b) A vector $\mathbf{x}$ belongs to $Y^\perp$ if and only if $\mathbf{x}$ is orthogonal to every vector in a set that spans Y.

Proof. See Exercises 4 and 6 on Coursework 9. □

We finish this section with an application of the concepts introduced so far to the column space and the nullspace of a matrix. These subspaces are sometimes called the **fundamental subspaces** of a matrix. The next theorem, a veritable gem of Linear Algebra, shows that the fundamental subspaces of a matrix and that of its transpose are intimately related via orthogonality:

Theorem 6.12 (Fundamental Subspace Theorem)**.** *Let $A \in \mathbb{R}^{m\times n}$. Then:*

(a) $N(A) = \mathrm{col}(A^T)^\perp$.

(b) $N(A^T) = \mathrm{col}(A)^\perp$.

Proof. In this proof we shall identify the rows of A (which are strictly speaking $1\times n$ matrices) with vectors in $\mathbb{R}^n$.

(a) Let $\mathbf{x} \in \mathbb{R}^n$. Then

$$\begin{aligned}\mathbf{x} \in N(A) &\iff A\mathbf{x} = \mathbf{0}\\ &\iff \mathbf{x} \text{ is orthogonal to every row of } A\\ &\iff \mathbf{x} \text{ is orthogonal to every column of } A^T\\ &\iff \mathbf{x} \in \mathrm{col}(A^T)^\perp\,,\end{aligned}$$

so $N(A) = \mathrm{col}(A^T)^\perp$.

(b) Apply (a) to A^T.

□

6.3 Orthogonal sets

In this section we shall investigate collections of vectors with the property that each vector is orthogonal to every other vector in the set. As we shall see, these sets have a number of astonishing properties.

Definition 6.13. A set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ in $\mathbb{R}^n$ is said to be an **orthogonal set** if each pair of distinct vectors is orthogonal, that is, if

$$\mathbf{u}_i \cdot \mathbf{u}_j = 0 \quad \text{whenever } i \neq j\,.$$

Example 6.14. If

$$\mathbf{u}_1 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} -1 \\ -4 \\ 7 \end{pmatrix},$$

then $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set since

$$\begin{aligned} \mathbf{u}_1 \cdot \mathbf{u}_2 &= 3 \cdot (-1) + 1 \cdot 2 + 1 \cdot 1 = 0 \\ \mathbf{u}_1 \cdot \mathbf{u}_3 &= 3 \cdot (-1) + 1 \cdot (-4) + 1 \cdot 7 = 0 \\ \mathbf{u}_2 \cdot \mathbf{u}_3 &= (-1) \cdot (-1) + 2 \cdot (-4) + 1 \cdot 7 = 0 \end{aligned}$$

The next theorem contains the first, perhaps surprising, property of orthogonal sets:

Theorem 6.15. *If $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is an orthogonal set of nonzero vectors, then the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_r$ are linearly independent.*

Proof. Suppose that

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_r\mathbf{u}_r = \mathbf{0}\,.$$

Then

$$\begin{aligned} 0 &= \mathbf{0} \cdot \mathbf{u}_1 \\ &= (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_r\mathbf{u}_r) \cdot \mathbf{u}_1 \\ &= c_1(\mathbf{u}_1 \cdot \mathbf{u}_1) + c_2(\mathbf{u}_2 \cdot \mathbf{u}_1) + \cdots + c_r(\mathbf{u}_r \cdot \mathbf{u}_1) \\ &= c_1(\mathbf{u}_1 \cdot \mathbf{u}_1)\,, \end{aligned}$$

since $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_2, \ldots, \mathbf{u}_r$. But since $\mathbf{u}_1$ is nonzero, $\mathbf{u}_1 \cdot \mathbf{u}_1$ is nonzero, so $c_1 = 0$. Similarly, $c_2, \ldots, c_r$ must be zero, and the assertion follows. □

Definition 6.16. An **orthogonal basis** for a subspace H of $\mathbb{R}^n$ is a basis of H that is also an orthogonal set.

The following theorem reveals why orthogonal bases are much 'nicer' than other bases in that the coordinates of a vector with respect to an orthogonal basis are spectacularly easy to compute:

Theorem 6.17. *Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ be an orthogonal basis for a subspace H of $\mathbb{R}^n$ and let $\mathbf{y} \in H$. If $c_1, \ldots, c_r$ are the coordinates of $\mathbf{y}$ with respect to $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$, that is,*

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_r\mathbf{u}_r\,,$$

then

$$c_j = \frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j} \quad \text{for each } j = 1, \ldots, r.$$

Proof. As in the proof of the preceding theorem, the orthogonality of $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ implies that

$$\mathbf{y} \cdot \mathbf{u}_1 = (c_1\mathbf{u}_1 + \cdots + c_r\mathbf{u}_r) \cdot \mathbf{u}_1 = c_1(\mathbf{u}_1 \cdot \mathbf{u}_1)\,.$$

Since $\mathbf{u}_1 \cdot \mathbf{u}_1$ is not zero (why?), we can solve for c_1 in the above equation and find the stated expression. In order to find c_j for $j = 2, \ldots, r$, compute $\mathbf{y} \cdot \mathbf{u}_j$ and solve for c_j. □

Example 6.18. Show that the set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ in Example 6.14 is an orthogonal basis for $\mathbb{R}^3$ and express the vector $\mathbf{y} = (6, 1, -8)^T$ as a linear combination of the basis vectors.

Solution. Note that by Theorem 6.15 the vectors in the orthogonal set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ are linearly independent, so must form a basis for $\mathbb{R}^3$, since $\dim \mathbb{R}^3 = 3$.

Now

$$\begin{aligned} \mathbf{y}\cdot\mathbf{u}_1 &= 11\,, & \mathbf{y}\cdot\mathbf{u}_2 &= -12\,, & \mathbf{y}\cdot\mathbf{u}_3 &= -66\,, \\ \mathbf{u}_1\cdot\mathbf{u}_1 &= 11\,, & \mathbf{u}_2\cdot\mathbf{u}_2 &= 6\,, & \mathbf{u}_3\cdot\mathbf{u}_3 &= 66\,, \end{aligned}$$

so

$$\mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 + \frac{\mathbf{y}\cdot\mathbf{u}_3}{\mathbf{u}_3\cdot\mathbf{u}_3}\mathbf{u}_3 = \frac{11}{11}\mathbf{u}_1 + \frac{-12}{6}\mathbf{u}_2 + \frac{-66}{66}\mathbf{u}_3 = \mathbf{u}_1 - 2\mathbf{u}_2 - \mathbf{u}_3\,.$$

□

6.4 Orthonormal sets

Definition 6.19. A set $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ of vectors in $\mathbb{R}^n$ is said to be an **orthonormal set** if it is an orthogonal set of unit vectors. Thus $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is an orthonormal set if and only if

$$\mathbf{u}_i\cdot\mathbf{u}_j = \delta_{ij} \quad \text{for } i, j = 1, \ldots, r\,,$$

where

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

If H is a subspace of $\mathbb{R}^n$ spanned by $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$, then $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is said to be an **orthonormal basis** for H.

Example 6.20. The standard basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ of $\mathbb{R}^n$ is an orthonormal set (and also an orthonormal basis for $\mathbb{R}^n$). Moreover, any nonempty subset of $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is an orthonormal set.

Here is a less trivial example:

Example 6.21. If

$$\mathbf{u}_1 = \begin{pmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix},$$

then $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal set, since

$$\begin{aligned} \mathbf{u}_1\cdot\mathbf{u}_2 &= -2/\sqrt{18} + 1/\sqrt{18} + 1/\sqrt{18} = 0 \\ \mathbf{u}_1\cdot\mathbf{u}_3 &= 0/\sqrt{12} - 1/\sqrt{12} + 1/\sqrt{12} = 0 \\ \mathbf{u}_2\cdot\mathbf{u}_3 &= 0/\sqrt{6} - 1/\sqrt{6} + 1/\sqrt{6} = 0 \end{aligned}$$

and

$$\begin{aligned} \mathbf{u}_1\cdot\mathbf{u}_1 &= 4/6 + 1/6 + 1/6 = 1 \\ \mathbf{u}_2\cdot\mathbf{u}_2 &= 1/3 + 1/3 + 1/3 = 1 \\ \mathbf{u}_3\cdot\mathbf{u}_3 &= 0/2 + 1/2 + 1/2 = 1 \end{aligned}$$

Moreover, since by Theorem 6.15 the vectors $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are linearly independent and $\dim \mathbb{R}^3 = 3$, the set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is a basis for $\mathbb{R}^3$. Thus $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal basis for $\mathbb{R}^3$.

Matrices whose columns form an orthonormal set are important in applications, in particular in computational algorithms. We are now going to explore some of their properties.

Theorem 6.22. *An $m \times n$ matrix U has orthonormal columns if and only if $U^TU = I$.*

Proof. As an illustration of the general idea, suppose for the moment that U has only three columns, each a vector in $\mathbb{R}^m$. Write

$$U = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{pmatrix} .$$

Then

$$U^TU = \begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1^T\mathbf{u}_1 & \mathbf{u}_1^T\mathbf{u}_2 & \mathbf{u}_1^T\mathbf{u}_3 \\ \mathbf{u}_2^T\mathbf{u}_1 & \mathbf{u}_2^T\mathbf{u}_2 & \mathbf{u}_2^T\mathbf{u}_3 \\ \mathbf{u}_3^T\mathbf{u}_1 & \mathbf{u}_3^T\mathbf{u}_2 & \mathbf{u}_3^T\mathbf{u}_3 \end{pmatrix}$$

so the (i, j)-entry of U^TU is just $\mathbf{u}_i \cdot \mathbf{u}_j$ and the assertion follows from the definition of orthonormality.

The proof for the general case is exactly the same, once you have convinced yourself that the (i, j)-entry of U^TU is the dot product of the i-th column of U with the j-th column of U. □

The following theorem is a simple consequence:

Theorem 6.23. *Let $U \in \mathbb{R}^{m\times n}$ have orthonormal columns, and let $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$. Then:*

(a) $(U\mathbf{x})\cdot(U\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$;

(b) $\|U\mathbf{x}\| = \|\mathbf{x}\|$;

(c) $(U\mathbf{x})\cdot(U\mathbf{y}) = 0$ *if and only if* $\mathbf{x}\cdot\mathbf{y} = 0$.

Proof. See Exercise 1 on Coursework 10. □

Here is a rephrasing of the theorem above in the language of linear transformations. Let $U \in \mathbb{R}^{m\times n}$ be a matrix with orthonormal columns and let L_U be the corresponding linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. Property (b) says that the mapping L_U preserves the lengths of vectors, and (c) says that L_U preserves orthogonality. These properties are important for many computer algorithms.

Before concluding this section we mention a class of matrices that fits naturally in the present context and which will play an important role in the next chapter:

Definition 6.24. A *square* matrix Q is said to be **orthogonal** if $Q^TQ = I$.

The above considerations show that every square matrix with orthonormal columns is an orthogonal matrix. Two other interesting properties of orthogonal matrices are contained in the following theorem.

Theorem 6.25. *Let $Q \in \mathbb{R}^{n\times n}$ be an orthogonal matrix. Then:*

(a) Q is invertible and $Q^{-1} = Q^T$;

(b) if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$, then $\{Q\mathbf{v}_1, \ldots, Q\mathbf{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$.

Proof. See Exercise 2 on Coursework 10. □

6.5 Orthogonal projections

In this section we shall study a particularly nice way of decomposing an arbitrary vector in $\mathbb{R}^n$. More precisely, if H is a subspace of $\mathbb{R}^n$ and $\mathbf{y}$ any vector in $\mathbb{R}^n$ then, as we shall see, we can write $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$ where $\hat{\mathbf{y}}$ is in H, and $\mathbf{z}$ is orthogonal to H. This is a very useful technique which has a number of interesting consequences, some of which you will see later in this chapter.

Theorem 6.26 (Orthogonal Decomposition Theorem)**.** *Let H be a subspace of $\mathbb{R}^n$. Then each $\mathbf{y}$ in $\mathbb{R}^n$ can be written uniquely in the form*

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}\,, \tag{6.1}$$

where $\hat{\mathbf{y}} \in H$ and $\mathbf{z} \in H^\perp$. In fact, if $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is an orthogonal basis for H, then

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}\cdot\mathbf{u}_r}{\mathbf{u}_r\cdot\mathbf{u}_r}\mathbf{u}_r\,, \tag{6.2}$$

and $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

Proof. Let $\hat{\mathbf{y}}$ be given by (6.2). Since $\hat{\mathbf{y}}$ is a linear combination of the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_r$, the vector $\hat{\mathbf{y}}$ must belong to H. Let $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$. Then

$$\begin{aligned}
\mathbf{z}\cdot\mathbf{u}_1 &= (\mathbf{y} - \hat{\mathbf{y}})\cdot\mathbf{u}_1 \\
&= \mathbf{y}\cdot\mathbf{u}_1 - \left(\frac{\mathbf{u}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\right)(\mathbf{u}_1\cdot\mathbf{u}_1) - 0 - \cdots - 0 \\
&= \mathbf{y}\cdot\mathbf{u}_1 - \mathbf{y}\cdot\mathbf{u}_1 \\
&= 0\,,
\end{aligned}$$

so $\mathbf{z}$ is orthogonal to $\mathbf{u}_1$. Similarly, we see that $\mathbf{z}$ is orthogonal to $\mathbf{u}_j$ for $j = 2, \ldots, r$, so $\mathbf{z} \in H^\perp$ by Theorem 6.11 (b).

In order to see that the decomposition (6.1) is unique, suppose that $\mathbf{y}$ can also be written as $\mathbf{y} = \hat{\mathbf{y}_1} + \mathbf{z}_1$, where $\hat{\mathbf{y}}_1 \in H$ and $\mathbf{z}_1 \in H^\perp$. Thus $\hat{\mathbf{y}} + \mathbf{z} = \hat{\mathbf{y}}_1 + \mathbf{z}_1$, so

$$\hat{\mathbf{y}} - \hat{\mathbf{y}}_1 = \mathbf{z}_1 - \mathbf{z}\,.$$

The above equality shows that the vector $\mathbf{v} = \hat{\mathbf{y}} - \hat{\mathbf{y}}_1$ belongs to both H and $H^\perp$. Thus $\mathbf{v}\cdot\mathbf{v} = 0$, which implies $\mathbf{v} = \mathbf{0}$. Therefore $\hat{\mathbf{y}} = \hat{\mathbf{y}}_1$ and $\mathbf{z} = \mathbf{z}_1$, so the decomposition (6.1) is unique. □

The vector $\hat{\mathbf{y}}$ in (6.1) is called the **orthogonal projection of $\mathbf{y}$ onto H** , and is written $\mathrm{proj}_H\mathbf{y}$, that is,

$$\hat{\mathbf{y}} = \mathrm{proj}_H\mathbf{y}\,.$$

One of the reasons why orthogonal projections play an important role in Linear Algebra, and indeed in other branches of Mathematics, is made plain in the following theorem:

Theorem 6.27 (Best Approximation Theorem)**.** *Let H be a subspace of $\mathbb{R}^n$, $\mathbf{y}$ any vector in $\mathbb{R}^n$, and $\hat{\mathbf{y}} = \mathrm{proj}_H\mathbf{y}$. Then $\hat{\mathbf{y}}$ is the closest point in H to $\mathbf{y}$, in the sense that*

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\| \tag{6.3}$$

for all $\mathbf{v} \in H$ distinct from $\hat{\mathbf{y}}$.

Proof. Take $\mathbf{v} \in H$ distinct from $\hat{\mathbf{y}}$. Then $\hat{\mathbf{y}} - \mathbf{v} \in H$. By the Orthogonal Decomposition Theorem, $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to H, so $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to $\hat{\mathbf{y}} - \mathbf{v}$.

Since

$$\mathbf{y} - \mathbf{v} = (\mathbf{y} - \hat{\mathbf{y}}) + (\hat{\mathbf{y}} - \mathbf{v}),$$

the Pythagorean Theorem (Theorem 6.8) gives

$$\|\mathbf{y} - \mathbf{v}\|^2 = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 + \|\hat{\mathbf{y}} - \mathbf{v}\|^2.$$

But $\|\hat{\mathbf{y}} - \mathbf{v}\|^2 > 0$, since $\hat{\mathbf{y}} \neq \mathbf{v}$, so the desired inequality (6.3) holds. □

The theorem above is the reason why the orthogonal projection of $\mathbf{y}$ onto H is often called the **best approximation of $\mathbf{y}$ by elements in H**.

We conclude this section with the following consequence of the Orthogonal Decomposition Theorem:

Theorem 6.28. *Let H be a subspace of $\mathbb{R}^n$. Then*

(a) $(H^\perp)^\perp = H$;

(b) $\dim H + \dim H^\perp = n$.

Proof. See Exercise 4 on Coursework 10. □

6.6 Gram Schmidt process

In the previous sections we saw on a number of occasions how useful orthogonal bases of subspaces can be. Witness, for example, the explicit expression for the orthogonal projection onto a subspace given in the Orthogonal Decomposition Theorem. So far we have not addressed the problem of how to manufacture an orthogonal basis. It turns out that there is a simple algorithm that does just that, namely producing an orthogonal basis for any nonzero subspace of $\mathbb{R}^n$:

Theorem 6.29 (Gram Schmidt process)**.** *Given a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$ of a subspace H of $\mathbb{R}^n$ define*

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1 \\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_r &= \mathbf{x}_r - \frac{\mathbf{x}_r \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_r \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \cdots - \frac{\mathbf{x}_r \cdot \mathbf{v}_{r-1}}{\mathbf{v}_{r-1} \cdot \mathbf{v}_{r-1}}\mathbf{v}_{r-1}
\end{aligned}$$

Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is an orthogonal basis for H. In addition

$$\operatorname{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \operatorname{Span}(\mathbf{x}_1, \ldots, \mathbf{x}_k) \quad \textit{for } 1 \leq k \leq r.$$

Proof. Write $H_k = \mathrm{Span}\,(\mathbf{x}_1, \ldots, \mathbf{x}_k)$. Set $\mathbf{v}_1 = \mathbf{x}_1$, so that $\mathrm{Span}\,(\mathbf{v}_1) = \mathrm{Span}\,(\mathbf{x}_1)$. Suppose that for some $k < r$ we have already constructed $\mathbf{v}_1, \ldots, \mathbf{v}_k$ so that $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is an orthogonal basis for H_k. Define

$$\mathbf{v}_{k+1} = \mathbf{x}_{k+1} - \mathrm{proj}_{H_k}\mathbf{x}_{k+1}\,.$$

By the Orthogonal Decomposition Theorem, $\mathbf{v}_{k+1}$ is orthogonal to H_k. Now the orthogonal projection $\mathrm{proj}_{H_k}\mathbf{x}_{k+1}$ belongs to H_k, which in turn is a subset of H_{k+1}, so $\mathbf{v}_{k+1} \in H_{k+1}$. Moreover, $\mathbf{v}_{k+1} \neq \mathbf{0}$, since $\mathbf{x}_{k+1} \notin H_k$. Thus $\{\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}\}$ is an orthogonal set of nonzero vectors in H_{k+1}. But $\dim H_{k+1} = k+1$, so $H_{k+1} = \mathrm{Span}\,(\mathbf{v}_1, \ldots, \mathbf{v}_{k+1})$. □

Remark 6.30. As with almost all the other results and techniques presented in this module, the best way to remember the Gram Schmidt process is to understand the proof. Here is the idea in a nut-shell: the Gram Schmidt process is an iterative procedure; if, at some stage, the orthogonal vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ have already been constructed, the next vector $\mathbf{v}_{k+1}$ is obtained by subtracting the orthogonal projection of $\mathbf{x}_{k+1}$ onto $H_k = \mathrm{Span}\,(\mathbf{v}_1, \ldots \mathbf{v}_k)$ from $\mathbf{x}_{k+1}$, that is,

$$\mathbf{v}_{k+1} = \mathbf{x}_{k+1} - \mathrm{proj}_{H_k}\mathbf{x}_{k+1}\,,$$

as this makes the vector $\mathbf{v}_{k+1}$ orthogonal H_k, and thus in particular, orthogonal to all previously constructed vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

Example 6.31. Let $H = \mathrm{Span}\,(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ where

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 6 \end{pmatrix}.$$

Clearly $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is a basis of H. Construct an orthogonal basis of H.

Solution. We start by setting

$$\mathbf{v}_1 = \mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

The vector $\mathbf{v}_2$ is constructed by subtracting the orthogonal projection of $\mathbf{x}_2$ onto $\mathrm{Span}\,(\mathbf{v}_1)$ from $\mathbf{x}_2$, that is,

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \mathbf{x}_2 - \frac{4}{4}\mathbf{v}_1 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

The vector $\mathbf{v}_3$ is constructed by subtracting the orthogonal projection of $\mathbf{x}_3$ onto $\mathrm{Span}\,(\mathbf{v}_1, \mathbf{v}_2)$ from $\mathbf{x}_3$, that is,

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = \mathbf{x}_3 - \frac{8}{4}\mathbf{v}_1 - \frac{6}{2}\mathbf{v}_2 = \begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \end{pmatrix},$$

producing the orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for H. □

6.7 Least squares problems

A type of problem that often arises in applications of Linear Algebra is to make sense of an overdetermined system

$$A\mathbf{x} = \mathbf{b}\,, \tag{6.4}$$

where $A \in \mathbb{R}^{m\times n}$ with $m > n$ and $\mathbf{b} \in \mathbb{R}^m$, for example the one I showed in class, which involved a 280×2 matrix. Clearly, the system (6.4) will not have a solution for every $\mathbf{b} \in \mathbb{R}^m$; in fact, as you will recall, the system has a solution if and only if $\mathbf{b} \in \mathrm{col}(A)$.

What do we do if we still demand a solution? The idea is to find an $\mathbf{x} \in \mathbb{R}^n$ that makes $A\mathbf{x}$ as close as possible to $\mathbf{b}$. In other words, in cases where no exact solution exists, we think of $A\mathbf{x}$ as an approximation to $\mathbf{b}$. The smaller the distance between $\mathbf{b}$ and $A\mathbf{x}$, given by $\|\mathbf{b} - A\mathbf{x}\|$, the better the approximation.

The general **least squares problem** is to find $\mathbf{x} \in \mathbb{R}^n$ that makes $\|\mathbf{b} - A\mathbf{x}\|$ as small as possible. Here, 'least squares' refers to the fact that $\|\mathbf{b} - A\mathbf{x}\|$ is the square root of a sum of squares.

Definition 6.32. Let $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. A **least squares solution of** $A\mathbf{x} = \mathbf{b}$ is an $\hat{\mathbf{x}} \in \mathbb{R}^n$ such that

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| \le \|\mathbf{b} - A\mathbf{x}\| \quad \text{for all } \mathbf{x} \in \mathbb{R}^n\,.$$

How do we find least squares solutions of a given system $A\mathbf{x} = \mathbf{b}$? To motivate the result to follow, let

$$\hat{\mathbf{b}} = \mathrm{proj}_{\mathrm{col}(A)}\mathbf{b}\,.$$

Since $\hat{\mathbf{b}}$ is in $\mathrm{col}(A)$, the equation $A\mathbf{x} = \hat{\mathbf{b}}$ is consistent, and there is an $\hat{\mathbf{x}} \in \mathbb{R}^n$ such that

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

By the Orthogonal Decomposition Theorem, $\mathbf{b} - \hat{\mathbf{b}}$ is orthogonal to $\mathrm{col}(A)$, so

$$\mathbf{b} - A\hat{\mathbf{x}} \in \mathrm{col}(A)^\perp\,.$$

But by the Fundamental Subspace Theorem $\mathrm{col}(A)^\perp = N(A^T)$, so

$$\mathbf{b} - A\hat{\mathbf{x}} \in N(A^T)\,.$$

Thus

$$A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}\,,$$

and hence

$$A^TA\hat{\mathbf{x}} = A^T\mathbf{b}\,.$$

To summarise what we have just said: a least squares solution of $A\mathbf{x} = \mathbf{b}$ satisfies

$$A^TA\mathbf{x} = A^T\mathbf{b}\,. \tag{6.5}$$

The matrix equation (6.5) represents a system of equations called the **normal equations** for $A\mathbf{x} = \mathbf{b}$.

Theorem 6.33. *Let $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. The set of least squares solutions of $A\mathbf{x} = \mathbf{b}$ coincides with the non-empty solution set of the normal equations*

$$A^TA\mathbf{x} = A^T\mathbf{b}\,. \tag{6.6}$$

Proof. We have just seen that a least squares solution $\hat{\mathbf{x}}$ must satisfy the normal equations. It turns out that the argument outlined also works in the reverse direction. To be precise, suppose that $\hat{\mathbf{x}}$ satisfies the normal equations, that is

$$A^T A\hat{\mathbf{x}} = A^T\mathbf{b}\,.$$

Then $A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}$, so $\mathbf{b} - A\hat{\mathbf{x}} \in N(A^T) = \mathrm{col}(A)^\perp$. Thus $\mathbf{b} - A\hat{\mathbf{x}}$ is orthogonal to $\mathrm{col}(A)$. Hence the equation

$$\mathbf{b} = A\hat{\mathbf{x}} + (\mathbf{b} - A\hat{\mathbf{x}})$$

is a decomposition of $\mathbf{b}$ into a sum of a vector in $\mathrm{col}(A)$ and a vector orthogonal to it. By the uniqueness of the orthogonal decomposition, $A\hat{\mathbf{x}}$ must be the orthogonal projection of $\mathbf{b}$ onto $\mathrm{col}(A)$. Thus, $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$, and $\hat{\mathbf{x}}$ is a least squares solution. □

Example 6.34. Find the least squares solution of the inconsistent system $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix}\,.$$

Solution. Compute

$$A^T A = \begin{pmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 17 & 1 \\ 1 & 5 \end{pmatrix}\,,$$

$$A^T\mathbf{b} = \begin{pmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix} = \begin{pmatrix} 19 \\ 11 \end{pmatrix}\,.$$

Thus the normal equations $A^T A\mathbf{x} = A^T\mathbf{b}$ are

$$\begin{pmatrix} 17 & 1 \\ 1 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 19 \\ 11 \end{pmatrix}\,.$$

This system can (and should, in general!) be solved by Gaussian elimination. In our case, however, it is quicker to spot that the coefficient matrix is invertible with inverse

$$(A^T A)^{-1} = \frac{1}{84}\begin{pmatrix} 5 & -1 \\ -1 & 17 \end{pmatrix}\,,$$

so $A^T A\mathbf{x} = A^T\mathbf{b}$ can be solved my multiplying both sides with $(A^T A)^{-1}$ from the left, giving the least squares solution

$$\hat{\mathbf{x}} = (A^T A)^{-1}A^T\mathbf{b} = \frac{1}{84}\begin{pmatrix} 5 & -1 \\ -1 & 17 \end{pmatrix}\begin{pmatrix} 19 \\ 11 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}\,.$$

□

Often (but not always!) the matrix $A^T A$ *is* invertible, and the method shown in the example above can be used. In general, the least squares solution need not be unique, and Gaussian elimination has to used to solve the normal equations. The following theorem gives necessary and sufficient conditions for $A^T A$ to be invertible.

Theorem 6.35. *Let $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. The matrix $A^T A$ is invertible if and only if the columns of A are linearly independent. In this case, $A\mathbf{x} = \mathbf{b}$ has only one least squares solution $\hat{\mathbf{x}}$, given by*

$$\hat{\mathbf{x}} = (A^T A)^{-1}A^T\mathbf{b}\,.$$

Proof. See Exercise 1 on Coursework 11 for the first part. The remaining assertion follows as in the previous example. □

Chapter 7

Eigenvalues and Eigenvectors

In this last chapter of our exploration of Linear Algebra we will revisit eigenvalues and eigenvectors of matrices, two concepts you have already encountered in Geometry I and, in various guises, in other modules. Let me first try to refresh your memory.

7.1 Definition and examples

If A is a square $n \times n$ matrix, we may regard it as a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$. This transformation sends a vector $\mathbf{x} \in \mathbb{R}^n$ to the vector $A\mathbf{x}$. For certain vectors, this action can be very simple.

Example 7.1. Let

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Then

$$A\mathbf{u} = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3\mathbf{u}$$
$$A\mathbf{w} = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 2\mathbf{w}$$

so the action of A on $\mathbf{u}$ and $\mathbf{w}$ is very easy to picture: it simply amounts to a stretching by 3 and 2, respectively.

Definition 7.2. An **eigenvector** of an $n \times n$ matrix A is a nonzero vector $\mathbf{x}$ such that

$$A\mathbf{x} = \lambda\mathbf{x}\,,$$

for some scalar λ. A scalar λ is called an **eigenvalue** of A if there is a non-trivial solution $\mathbf{x}$ to $A\mathbf{x} = \lambda\mathbf{x}$, in which case we say that $\mathbf{x}$ is an **eigenvector corresponding to the eigenvalue** λ.

Remark 7.3. Note that if $\mathbf{x}$ is an eigenvector of a matrix A with eigenvalue λ, then any nonzero multiple of $\mathbf{x}$ is also an eigenvector corresponding to λ, since

$$A(\alpha\mathbf{x}) = \alpha A\mathbf{x} = \alpha\lambda\mathbf{x} = \lambda(\alpha\mathbf{x})\,.$$

In the example above, $\mathbf{u}$ is an eigenvector of A corresponding to the eigenvalue 3, and $\mathbf{w}$ is an eigenvector of A corresponding to the eigenvalue 2. One of the reasons why eigenvalues and eigenvectors play a prominent role in Linear Algebra is that knowledge of the eigenvalues and eigenvectors of a matrix sometimes goes a long way to understand the action of A on *any* vector. In order to see this, suppose that $\mathbf{x}$ is an arbitrary vector in $\mathbb{R}^2$. Since $\mathbf{u}$ and $\mathbf{w}$ are linearly independent, they form a basis of $\mathbb{R}^2$, and we can write

$$\mathbf{x} = c_1\mathbf{u} + c_2\mathbf{w}$$

for some scalars c_1 and c_2. But then

$$A\mathbf{x} = c_1A\mathbf{u} + c_2A\mathbf{w} = 3c_1\mathbf{u} + 2c_2\mathbf{w}\,,$$

so if we define the basis $\mathcal{B} = \{\mathbf{u}, \mathbf{w}\}$ of $\mathbb{R}^2$, then

$$[A\mathbf{x}]_{\mathcal{B}} = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} [\mathbf{x}]_{\mathcal{B}}\,.$$

In other words, relative to the basis $\mathcal{B}$, the action of A on an arbitrary vector is very easy to understand. We shall explore this observation in more detail shortly, and we shall see that it holds the key to a new and powerful perspective on matrices.

We shall now investigate how to determine all the eigenvalues and eigenvectors of an $n \times n$ matrix A. We start by observing that the defining equation $A\mathbf{x} = \lambda\mathbf{x}$ can be written

$$(A - \lambda I)\mathbf{x} = \mathbf{0}\,. \tag{7.1}$$

Thus λ is an eigenvalue of A if and only if (7.1) has a non-trivial solution. The set of solutions of (7.1) is $N(A - \lambda I)$, that is, the nullspace of $A - \lambda I$, which is a subspace of $\mathbb{R}^n$.Thus, λ is an eigenvalue of A if and only if

$$N(A - \lambda I) \neq \{\mathbf{0}\}\,,$$

and any nonzero vector in $N(A - \lambda I)$ is an eigenvector belonging to λ. Moreover, by the Invertible Matrix Theorem, (7.1) has a non-trivial solution if and only if the matrix $A - \lambda I$ is singular, or equivalently

$$\det(A - \lambda I) = 0\,. \tag{7.2}$$

Notice now that if the determinant in (7.2) is expanded we obtain a polynomial of degree n in the variable λ,

$$p(\lambda) = \det(A - \lambda I)\,,$$

called the **characteristic polynomial of** A , and equation (7.2) is called the **characteristic equation of** A. So, in other words, the roots of the characteristic polynomial of A are exactly the eigenvalues of A. The following theorem summarises our findings so far:

Theorem 7.4. *Let A be an $n \times n$ matrix and λ a scalar. The following statements are equivalent:*

(a) λ is an eigenvalue of A;

(b) $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has a non-trivial solution;

(c) $N(A - \lambda I) \neq \{\mathbf{0}\}$;

(d) $A - \lambda I$ *is singular;*

(e) $\det(A - \lambda I) = 0$.

In view of the above theorem the following concept arises naturally:

Definition 7.5. If A is a square matrix and λ an eigenvalue of A, then $N(A - \lambda I)$ is called the **eigenspace corresponding to** λ.

Note that if λ is an eigenvalue of a matrix A, then every nonzero vector in the corresponding eigenspace $N(A - \lambda I)$ is an eigenvector corresponding to λ, and conversely, the set of all eigenvectors corresponding to λ together with the zero vector forms the corresponding eigenspace $N(A - \lambda I)$.

We shall now see how to use (e) in the theorem above to determine the eigenvalues and eigenvectors of a given matrix.

Example 7.6. Find all eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} -7 & -6 \\ 9 & 8 \end{pmatrix}.$$

Proof. First we calculate the characteristic polynomial of A

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} -7-\lambda & -6 \\ 9 & 8-\lambda \end{vmatrix} = (-7-\lambda)(8-\lambda) - (-6)\cdot 9 \\ &= -56 + 7\lambda - 8\lambda + \lambda^2 + 54 = \lambda^2 - \lambda - 2 = (\lambda+1)(\lambda-2)\,. \end{aligned}$$

Thus the characteristic equation is

$$(\lambda+1)(\lambda-2) = 0\,,$$

so the eigenvalues of the matrix are $\lambda_1 = -1$ and $\lambda_2 = 2$.

In order to find the eigenvectors belonging to $\lambda_1 = -1$ we must determine the nullspace of $A - \lambda_1 I = A + I$. Or, put differently, we need to determine all solutions of the system $(A+I)\mathbf{x} = \mathbf{0}$. This can be done using your favourite method, but, for reasons which will become clear in the next example, I strongly recommend Gaussian elimination, that is, we bring the augmented matrix $(A + I|\mathbf{0})$ to row echelon form:

$$(A+I|\mathbf{0}) = \left(\begin{array}{cc|c} -7+1 & -6 & 0 \\ 9 & 8+1 & 0 \end{array}\right) = \left(\begin{array}{cc|c} -6 & -6 & 0 \\ 9 & 9 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 9 & 9 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right),$$

so setting $x_2 = \alpha$ we find $x_1 = -x_2 = -\alpha$. Thus every vector in $N(A+I)$ is of the form

$$\begin{pmatrix} -\alpha \\ \alpha \end{pmatrix} = \alpha \begin{pmatrix} -1 \\ 1 \end{pmatrix},$$

so the eigenspace corresponding to the eigenvalue -1 is

$$\left\{ \alpha \begin{pmatrix} -1 \\ 1 \end{pmatrix} \,\middle|\, \alpha \in \mathbb{R} \right\},$$

and any nonzero multiple of $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ is an eigenvector corresponding to the eigenvalue -1.

Similarly, in order to find the eigenvectors belonging to $\lambda_2 = 2$ we bring $(A - \lambda_2 I|\mathbf{0}) = (A - 2I|\mathbf{0})$ to row echelon form:

$$(A - 2I|\mathbf{0}) = \left(\begin{array}{cc|c} -7-2 & -6 & 0 \\ 9 & 8-2 & 0 \end{array}\right) = \left(\begin{array}{cc|c} -9 & -6 & 0 \\ 9 & 6 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & \frac{2}{3} & 0 \\ 9 & 6 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & \frac{2}{3} & 0 \\ 0 & 0 & 0 \end{array}\right),$$

so setting $x_2 = \alpha$ we find $x_1 = -\frac{2}{3}x_2 = -\frac{2}{3}\alpha$. Thus every vector in $N(A-2I)$ is of the form

$$\begin{pmatrix} -\frac{2}{3}\alpha \\ \alpha \end{pmatrix} = \frac{\alpha}{3}\begin{pmatrix} -2 \\ 3 \end{pmatrix},$$

so the eigenspace corresponding to the eigenvalue 2 is

$$\left\{ \alpha \begin{pmatrix} -2 \\ 3 \end{pmatrix} \,\middle|\, \alpha \in \mathbb{R} \right\},$$

and any nonzero multiple of $\begin{pmatrix} -2 \\ 3 \end{pmatrix}$ is an eigenvector corresponding to the eigenvalue 2. □

Before we continue with another example you might want to have another look at the above calculations of eigenspaces. Observe that since we need to solve a *homogeneous* linear system there is no need to write down the right-most column of the augmented matrix (since it consists only of zeros); we simply perform elementary row operations on the coefficient matrix, keeping in mind that the right-most column of the augmented matrix will remain the zero column. We shall use this short-cut in all the following calculations of eigenspaces.

Example 7.7. Let

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

Find the eigenvalues and corresponding eigenspaces.

Solution. A slightly tedious calculation using repeated cofactor expansions shows that the characteristic polynomial of A is

$$\det(A - \lambda I) = \begin{vmatrix} 2-\lambda & -3 & 1 \\ 1 & -2-\lambda & 1 \\ 1 & -3 & 2-\lambda \end{vmatrix} = -\lambda(\lambda - 1)^2,$$

so the eigenvalues of A are $\lambda_1 = 0$ and $\lambda_2 = 1$.

In order to find the eigenspace corresponding to λ_1 we find the nullspace of $A - \lambda_1 I = A$ using Gaussian elimination:

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

so setting $x_3 = \alpha$ we find $x_2 = 0 - (-1)x_3 = \alpha$ and $x_1 = 0 - (-1)x_3 = \alpha$. Thus, every vector in $N(A)$ is of the form

$$\begin{pmatrix} \alpha \\ \alpha \\ \alpha \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

so the eigenspace corresponding to the eigenvalue 0 is

$$\left\{ \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \middle| \, \alpha \in \mathbb{R} \right\} .$$

In order to find the eigenspace corresponding to λ_2 we find the nullspace of $A-\lambda_2 I = A-I$, again using Gaussian elimination:

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

so setting $x_2 = \alpha$ and $x_3 = \beta$ we find $x_1 = 3x_2 - x_3 = 3\alpha - \beta$. Thus every vector in $N(A-I)$ is of the form

$$\begin{pmatrix} 3\alpha - \beta \\ \alpha \\ \beta \end{pmatrix} = \alpha \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} ,$$

and so the eigenspace corresponding to the eigenvalue 1 is

$$\left\{ \alpha \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \middle| \, \alpha, \beta \in \mathbb{R} \right\} .$$

□

Example 7.8. Find the eigenvalues of the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} .$$

Solution. Using the fact that the determinant of a triangular matrix is the product of the diagonal entries we find

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 2 & 3 \\ 0 & 4-\lambda & 5 \\ 0 & 0 & 6-\lambda \end{vmatrix} = (1-\lambda)(4-\lambda)(6-\lambda) ,$$

so the eigenvalues of A are 1, 4, and 6. □

The above example and its method of solution are easily generalised:

Theorem 7.9. *The eigenvalues of a triangular matrix are precisely the diagonal entries of the matrix.*

The next theorem gives an important sufficient (but not necessary) condition for two matrices to have the same eigenvalues. It also serves as the foundation for many numerical procedures to approximate eigenvalues of matrices, some of which you will encounter if you take the module MTH5110, Introduction to Numerical Computing.

Theorem 7.10. *Let A and B be two $n \times n$ matrices and suppose that A and B are similar, that is, there is an invertible matrix $S \in \mathbb{R}^{n \times n}$ such that $B = S^{-1}AS$. Then A and B have the same characteristic polynomial, and, consequently, have the same eigenvalues.*

Proof. If $B = S^{-1}AS$, then

$$B - \lambda I = S^{-1}AS - \lambda I = S^{-1}AS - \lambda S^{-1}S = S^{-1}(AS - \lambda S) = S^{-1}(A - \lambda I)S\,.$$

Thus, using the multiplicativity of determinants,

$$\det(B - \lambda I) = \det(S^{-1})\det(A - \lambda I)\det(S) = \det(A - \lambda I)\,,$$

because $\det(S^{-1})\det(S) = \det(S^{-1}S) = \det(I) = 1$. □

We will revisit this theorem from a different perspective in the next section.

7.2 Diagonalisation

In many applications of Linear Algebra one is faced with the following problem: given a square matrix A, find the k-th power A^k of A for large values of k. In general, this can be a very onerous task. For certain matrices, however, evaluating powers is spectacularly easy:

Example 7.11. Let $D \in \mathbb{R}^{2\times 2}$ be given by

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

Then

$$D^2 = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 2^2 & 0 \\ 0 & 3^2 \end{pmatrix}$$

and

$$D^3 = DD^2 = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 2^2 & 0 \\ 0 & 3^2 \end{pmatrix} = \begin{pmatrix} 2^3 & 0 \\ 0 & 3^3 \end{pmatrix}.$$

In general,

$$D^k = \begin{pmatrix} 2^k & 0 \\ 0 & 3^k \end{pmatrix},$$

for $k \geq 1$.

After having had another look at the example above, you should convince yourself that if D is a *diagonal* $n \times n$ matrix with diagonal entries $d_1, \ldots, d_n$, then D^k is the diagonal matrix whose diagonal entries are $d_1^k, \ldots, d_n^k$. The moral of this is that calculating powers for diagonal matrices is easy. What if the matrix is not diagonal? The next best situation arises if the matrix is *similar* to a diagonal matrix. In this case, calculating powers is almost as easy as calculating powers of diagonal matrices, as we shall see shortly. We shall now single out matrices with this property and give them a special name:

Definition 7.12. An $n \times n$ matrix A is said to be **diagonalisable** if it is similar to a diagonal matrix, that is, if there is an invertible matrix $P \in \mathbb{R}^{n\times n}$ such that

$$P^{-1}AP = D\,,$$

where D is a diagonal matrix. In this case we say that P **diagonalises** A.

Note that if A is a matrix which is diagonalised by P, that is, $P^{-1}AP = D$ with D diagonal, then

$$\begin{aligned} A &= PDP^{-1}\,, \\ A^2 &= PDP^{-1}PDP^{-1} = PD^2P^{-1}\,, \\ A^3 &= AA^2 = PDP^{-1}PD^2P^{-1} = PD^3P^{-1}\,, \end{aligned}$$

and in general

$$A^k = PD^kP^{-1}\,,$$

for any $k \geq 1$. Thus powers of A are easily computed, as claimed.

Before giving a characterisation of diagonalisable matrices we require an auxiliary result of independent interest:

Theorem 7.13. *If $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix A, then the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly independent.*

Proof. By contradiction. Suppose to the contrary that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly dependent. We may then assume (after reordering the $\mathbf{v}$'s and the λ's if necessary), that there is an index $p < r$ such that $\mathbf{v}_{p+1}$ is a linear combination of the preceding linearly independent vectors. Thus there exist scalars $c_1, \ldots, c_p$ such that

$$c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{v}_{p+1}\,. \tag{7.3}$$

Multiplying both sides of the above equation by A and using the fact that $A\mathbf{v}_k = \lambda_k\mathbf{v}_k$ for each k, we obtain

$$c_1\lambda_1\mathbf{v}_1 + \cdots + c_p\lambda_p\mathbf{v}_p = \lambda_{p+1}\mathbf{v}_{p+1}\,. \tag{7.4}$$

Multiplying both sides of (7.3) by λ_{p+1} and subtracting the result from (7.4), we see that

$$c_1(\lambda_1 - \lambda_{p+1})\mathbf{v}_1 + \cdots + c_p(\lambda_p - \lambda_{p+1})\mathbf{v}_p = \mathbf{0}\,. \tag{7.5}$$

Since the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are linearly independent, the coefficients in (7.5) must all be zero. But none of the factors $\lambda_k - \lambda_{p+1}$ is zero, because the eigenvalues are distinct, so we must have $c_1 = \ldots = c_p = 0$. But then (7.3) implies that $\mathbf{v}_{p+1} = \mathbf{0}$, which is impossible, because $\mathbf{v}_{p+1}$ is an eigenvector. Thus our assumption that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly dependent must be false, that is, the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly independent. □

We are now able to prove the main the result of this section:

Theorem 7.14 (Diagonalisation Theorem)**.** *An $n \times n$ matrix A is diagonalisable if and only if A has n linearly independent eigenvectors.*

In fact, $P^{-1}AP = D$, with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A. In this case, the diagonal entries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in P.

Proof. First observe that if P is any $n \times n$ matrix with columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$, and if D is any diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$, then

$$AP = A\begin{pmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{pmatrix} = \begin{pmatrix}A\mathbf{v}_1 & \cdots & A\mathbf{v}_n\end{pmatrix}\,, \tag{7.6}$$

while

$$PD = \begin{pmatrix}\lambda\mathbf{v}_1 & \cdots & \lambda_n\mathbf{v}_n\end{pmatrix}\,. \tag{7.7}$$

First we prove the 'only if' part of the theorem. Suppose that A is diagonalisable and that $P^{-1}AP = D$. Then $AP = PD$, and equations (7.6) and (7.7) imply that

$$\begin{pmatrix} A\mathbf{v}_1 & \cdots & A\mathbf{v}_n \end{pmatrix} = \begin{pmatrix} \lambda_1\mathbf{v}_1 & \cdots & \lambda_n A\mathbf{v}_n \end{pmatrix} . \tag{7.8}$$

Thus

$$A\mathbf{v}_1 = \lambda_1\mathbf{v}_1 , \quad A\mathbf{v}_2 = \lambda_2\mathbf{v}_2 , \quad \ldots , \quad A\mathbf{v}_n = \lambda_n\mathbf{v}_n . \tag{7.9}$$

Since P is invertible its columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$ must be linearly independent. Moreover, none of its columns can be zero, so (7.9) implies that $\lambda_1, \ldots, \lambda_n$ are eigenvalues of A with corresponding eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. This finishes the proof of the 'only if' part.

In order to prove the 'if' part, suppose that A has n linearly independent eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$. Let P be the matrix whose columns are $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and let D be the diagonal matrix whose diagonal entries are $\lambda_1, \ldots, \lambda_n$. Then equations (7.6), (7.7), and (7.8) imply that $AP = PD$. But since the columns of P are linearly independent, P is invertible, so $P^{-1}AP = D$ as claimed. □

Example 7.15. Diagonalise the following matrix, if possible:

$$A = \begin{pmatrix} -7 & 3 & -3 \\ -9 & 5 & -3 \\ 9 & -3 & 5 \end{pmatrix} .$$

Solution. A slightly tedious calculation shows that the characteristic polynomial is given by

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} -7-\lambda & 3 & -3 \\ -9 & 5-\lambda & -3 \\ 9 & -3 & 5-\lambda \end{vmatrix} = -\lambda^3 + 3\lambda^2 - 4 .$$

The cubic p above can be factored by spotting that -1 is a root. Polynomial division then yields

$$p(\lambda) = -(\lambda+1)(\lambda^2 - 4\lambda + 4) = -(\lambda+1)(\lambda-2)^2 ,$$

so the distinct eigenvalues of A are 2 and -1.

The usual methods (see Examples 7.6 and 7.7) now produce a basis for each of the two eigenspaces and it turns out that

$$N(A - 2I) = \mathrm{Span}\,(\mathbf{v}_1, \mathbf{v}_2), \quad \text{where} \quad \mathbf{v}_1 = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} , \quad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix} ,$$

$$N(A + I) = \mathrm{Span}\,(\mathbf{v}_3), \quad \text{where} \quad \mathbf{v}_3 = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} .$$

You may now want to confirm, using your favourite method, that the three vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent. As we shall see shortly, this is not really necessary: the union of basis vectors for eigenspaces always produces linearly independent vectors (see the proof of Corollary 7.17 below).

Thus, A is diagonalisable, since it has 3 linearly independent eigenvectors. In order to find the diagonalising matrix P we recall that defining

$$P = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & -1 \\ 3 & 0 & -1 \\ 0 & 3 & 1 \end{pmatrix}$$

does the trick, that is, $P^{-1}AP = D$, where D is the diagonal matrix whose entries are the eigenvalues of A and where the order of the eigenvalues matches the order chosen for the eigenvectors in P, that is,

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

It is good practice to check that P and D really do the job they are supposed to do:

$$AP = \begin{pmatrix} -7 & 3 & -3 \\ -9 & 5 & -3 \\ 9 & -3 & 5 \end{pmatrix} \begin{pmatrix} 1 & -1 & -1 \\ 3 & 0 & -1 \\ 0 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 1 \\ 6 & 0 & 1 \\ 0 & 6 & -1 \end{pmatrix} ,$$

$$PD = \begin{pmatrix} 1 & -1 & -1 \\ 3 & 0 & -1 \\ 0 & 3 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 1 \\ 6 & 0 & 1 \\ 0 & 6 & -1 \end{pmatrix} ,$$

so $AP = PD$, and hence $P^{-1}AP = D$ as required. □

Example 7.16. Diagonalise the following matrix, if possible:

$$A = \begin{pmatrix} -6 & 3 & -2 \\ -7 & 5 & -1 \\ 8 & -3 & 4 \end{pmatrix} .$$

Solution. The characteristic polynomial of A turns out to be exactly the same as in the previous example:

$$\det(A - \lambda I) = -\lambda^3 + 3\lambda^2 - 4 = -(\lambda + 1)(\lambda - 2)^2 .$$

Thus the eigenvalues of A are 2 and -1. However, in this case it turns out that both eigenspaces are 1-dimensional:

$$N(A - 2I) = \mathrm{Span}\,(\mathbf{v}_1) \quad \text{where} \quad \mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} ,$$

$$N(A + I) = \mathrm{Span}\,(\mathbf{v}_2) \quad \text{where} \quad \mathbf{v}_1 = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} .$$

Since A has only 2 linearly independent eigenvectors, the Diagonalisation Theorem implies that A is not diagonalisable. □

Put differently, the Diagonalisation Theorem states that a matrix $A \in \mathbb{R}^{n\times n}$ is diagonalisable if and only if A has enough eigenvectors to form a basis of $\mathbb{R}^n$. The following corollary makes this restatement even more precise:

Corollary 7.17. *Let $A \in \mathbb{R}^{n\times n}$ and let $\lambda_1, \ldots, \lambda_r$ be the (distinct) eigenvalues of A. Then A is diagonalisable if and only if*

$$\dim N(A - \lambda_1) + \cdots + \dim N(A - \lambda_r I) = n .$$

Proof. For simplicity we will only consider the case where A has two distinct eigenvalues; the general case can be proved along the same lines. Suppose that A has two distinct eigenvalues λ_1 and λ_2. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ be a basis for $N(A - \lambda_1 I)$ and let $\{\mathbf{w}_1, \ldots, \mathbf{w}_q\}$ be a basis for $N(A - \lambda_2 I)$. First we show that the $p + q$ vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p, \mathbf{w}_1, \ldots, \mathbf{w}_q$ are linearly independent. In order to see this suppose that there are scalars $\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q$ such that

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_p \mathbf{v}_p + \beta_1 \mathbf{w}_1 + \cdots + \beta_q \mathbf{w}_q = \mathbf{0}\,. \tag{7.10}$$

Theorem 7.13 now implies that

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_p \mathbf{v}_p = \mathbf{0} \quad \text{and} \quad \beta_1 \mathbf{w}_1 + \cdots + \beta_q \mathbf{w}_q = \mathbf{0}\,, \tag{7.11}$$

because otherwise (7.10) would state that $\mathbf{0}$ can be written as a nontrivial linear combination of eigenvectors belonging to distinct eigenvalues of A. But since $\mathbf{v}_1, \ldots, \mathbf{v}_p$ and $\mathbf{w}_1, \ldots \mathbf{w}_q$ are linearly independent, it follows that $\alpha_1 = \cdots = \alpha_p = \beta_1 = \cdots = \beta_q = 0$. Thus the $p + q$ vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p, \mathbf{w}_1, \ldots, \mathbf{w}_q$ are linearly independent as claimed.

By the Diagonalisation Theorem A is diagonalisable if and only if A has n linearly independent eigenvectors. By what we have just proved, this is the case if and only if $p + q = n$, and the corollary follows because $\dim N(A - \lambda_1) = p$ and $\dim N(A - \lambda_2 I) = q$. □

A very useful special case of the Diagonalisation Theorem is the following:

Theorem 7.18. *An $n \times n$ matrix with n distinct eigenvalues is diagonalisable.*

Proof. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be eigenvectors corresponding to the n distinct eigenvalues of A. Then the n vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent by Theorem 7.13. Hence A is diagonalisable by the Diagonalisation Theorem. □

Remark 7.19. Note that the above condition for diagonalisability is *sufficient* but not *necessary*: an $n \times n$ matrix which does not have n distinct eigenvalues may or may not be diagonalisable (see Examples 7.15 and 7.16).

Example 7.20. The matrix

$$A = \begin{pmatrix} 1 & -1 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{pmatrix}$$

is diagonalisable, since it has three distinct eigenvalues 1, 2, and 3.

7.3 Interlude: complex vector spaces and matrices

Consider the matrix

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

What are the eigenvalues of A? Notice that

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1\,,$$

so the characteristic polynomial does not have any real roots, and hence A does not have any real eigenvalues. However, since

$$\lambda^2 + 1 = \lambda^2 - (-1) = \lambda - i^2 = (\lambda - i)(\lambda + i)\,,$$

the characteristic polynomial has two *complex* roots, namely i and $-i$. Thus it makes sense to say that A has two complex eigenvalues i and $-i$. What are the corresponding eigenvectors? Solving

$$(A - iI)\mathbf{x} = \mathbf{0}$$

leads to the system

$$\begin{array}{rcrcl} -ix_1 & - & x_2 & = & 0 \\ x_1 & + & ix_2 & = & 0 \end{array}$$

Both equations yield the condition $x_2 = -ix_1$, so $\begin{pmatrix} 1 \\ -i \end{pmatrix}$ is an eigenvector corresponding to the eigenvalue i. Indeed

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \begin{pmatrix} i \\ 1 \end{pmatrix} = \begin{pmatrix} i \\ -i^2 \end{pmatrix} = i \begin{pmatrix} 1 \\ -i \end{pmatrix} .$$

Similarly, we see that $\begin{pmatrix} 1 \\ i \end{pmatrix}$ is an eigenvector corresponding to the eigenvalue $-i$. Indeed

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} = \begin{pmatrix} -i \\ 1 \end{pmatrix} = \begin{pmatrix} -i \\ -i^2 \end{pmatrix} = -i \begin{pmatrix} 1 \\ i \end{pmatrix} .$$

The moral of this example is the following: on the one hand, we could just say that the matrix A has no real eigenvalues and stop the discussion right here. On the other hand, we just saw that it makes sense to say that A has two complex eigenvalues with corresponding complex eigenvectors.

This leads to the idea of leaving our current real set-up, and enter a complex realm instead. As it turns out, this is an immensely powerful idea. However, as our time is limited, we shall only cover the bare necessities, allowing us to prove the main result of the next section.

Let $\mathbb{C}^n$ denote the set of all n-vectors with complex entries, that is,

$$\mathbb{C}^n = \left\{ \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \;\middle|\; z_1, \ldots, z_n \in \mathbb{C} \right\} .$$

Just as in $\mathbb{R}^n$, we add vectors in $\mathbb{C}^n$ by adding their entries, and we can multiply a vector in $\mathbb{C}^n$ by a complex number, by multiplying each entry.

Example 7.21. Let $\mathbf{z}, \mathbf{w} \in \mathbb{C}^3$ and $\alpha \in \mathbb{C}$, with

$$\mathbf{z} = \begin{pmatrix} 1+i \\ 2i \\ 3 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} -2+3i \\ 1 \\ 2+i \end{pmatrix}, \quad \alpha = (1+2i) .$$

Then

$$\mathbf{z} + \mathbf{w} = \begin{pmatrix} (1+i)+(-2+3i) \\ 2i+1 \\ 3+(2+i) \end{pmatrix} = \begin{pmatrix} -1+4i \\ 1+2i \\ 5+i \end{pmatrix}$$

$$\alpha\mathbf{z} = \begin{pmatrix} (1+2i)(1+i) \\ (1+2i)(2i) \\ (1+2i)\cdot 3 \end{pmatrix} = \begin{pmatrix} 1+2i+i+2i^2 \\ 2i+(2i)^2 \\ 3+6i \end{pmatrix} = \begin{pmatrix} -1+3i \\ -4+2i \\ 3+6i \end{pmatrix}$$

If addition and scalar multiplication is defined in this way (now allowing scalars to be in $\mathbb{C}$), then $\mathbb{C}^n$ satisfies all the axioms of a vector space. Similarly, we can introduce the set of all $m \times n$ matrices with complex entries, call it $\mathbb{C}^{m\times n}$, and define addition and scalar multiplication (again allowing complex scalars) entry-wise just as in $\mathbb{R}^{m\times n}$. Again, $\mathbb{C}^{m\times n}$ satisfies all the axioms of a vector space.

Fact 7.22. *All the results in Chapters 1–5, and all the results from the beginning of this chapter hold verbatim, if 'scalar' is taken to mean 'complex number'.*

Since 'scalars' are now allowed to be complex numbers, $\mathbb{C}^n$ and $\mathbb{C}^{m\times n}$ are known as **complex vector spaces**.

The reason for allowing this more general set-up is that, in a certain sense, complex numbers are much nicer than real numbers. More precisely, we have the following result:

Theorem 7.23 (Fundamental Theorem of Algebra)**.** *If p is a complex polynomial of degree $n \geq 1$, that is,*

$$p(z) = c_n z^n + \cdots + c_1 z + c_0\,,$$

where $c_0, c_1, \ldots, c_n \in \mathbb{C}$, then p has at least one (possibly complex) root.

Corollary 7.24. *Every matrix $A \in \mathbb{C}^{n\times n}$ has at least one (possibly complex) eigenvalue and a corresponding eigenvector $\mathbf{z} \in \mathbb{C}^n$.*

Proof. Since λ is an eigenvalue of A if and only if $\det(A - \lambda I) = 0$ and since $p(\lambda) = \det(A - \lambda I)$ is a polynomial with complex coefficients of degree n, the assertion follows from the Fundamental Theorem of Algebra. □

The corollary above is the main reason why complex vector spaces are considered. We are *guaranteed* that every matrix has at least one eigenvalue, and we may then use the powerful tools developed in the earlier parts of this chapter to analyse matrices through their eigenvalues and eigenvectors.

7.4 Spectral Theorem for Symmetric Matrices

This last section of the last chapter is devoted to one of the gems of Linear Algebra: the Spectral Theorem. This result, which has many applications, a number of which you will see in other modules, is concerned with the diagonalisability of *symmetric* matrices. Recall that a matrix $A \in \mathbb{R}^{n\times n}$ is said to be diagonalisable if there is an invertible matrix $P \in \mathbb{R}^{n\times n}$ such that $P^{-1}AP$ is diagonal. We already know that A is diagonalisable if and only if A has n linearly independent eigenvectors. However, this condition is difficult to check in practice. It may thus come as a surprise that there is a sufficiently rich class of matrices that are always diagonalisable, and moreover, that the diagonalising matrix P is of a special form. This is the content of the Spectral Theorem for Symmetric Matrices, or Spectral Theorem for short:[1]

Theorem 7.25 (Spectral Theorem for Symmetric Matrices)**.** *Let $A \in \mathbb{R}^{n\times n}$ be symmetric. Then there is an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ such that*

$$Q^T A Q = D,$$

where $D \in \mathbb{R}^{n\times n}$ is diagonal.

Or put differently: every symmetric matrix can be diagonalised by an orthogonal matrix.

[1] There are other, more general versions of the Spectral Theorem. In this course, we will only consider the symmetric case.

The proof of this theorem requires a number of auxiliary results which we shall first state and prove.

Lemma 7.26. *The product of two orthogonal matrices of the same size is orthogonal, that is, if $Q_0 \in \mathbb{R}^{n\times n}$ and $Q_1 \in \mathbb{R}^{n\times n}$ are orthogonal, then Q_0Q_1 is also orthogonal.*

Proof. Since, by definition, $Q_0^TQ_0 = I$ and $Q_1^TQ_1 = I$, we have

$$(Q_0Q_1)^TQ_0Q_1 = Q_1^TQ_0^TQ_0Q_1 = Q_1^TQ_1 = I\,,$$

so Q_0Q_1 is orthogonal. □

Lemma 7.27. *Let $A \in \mathbb{R}^{n\times n}$ be symmetric. Then, for any vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,*

$$(A\mathbf{x})\cdot\mathbf{y} = \mathbf{x}\cdot(A\mathbf{y})\,.$$

Proof. Since $A^T = A$, we have

$$(A\mathbf{x})\cdot\mathbf{y} = (A\mathbf{x})^T\mathbf{y} = \mathbf{x}^TA^T\mathbf{y} = \mathbf{x}^TA\mathbf{y} = \mathbf{x}\cdot(A\mathbf{y})\,.$$

□

Lemma 7.28. *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix. If $\mathbf{v} \in \mathbb{R}^n$ is an eigenvector of A and $H = \mathrm{Span}\,(\mathbf{v})$, then*

$$L_A(H^\perp) \subset H^\perp\,,$$

that is,

$$\mathbf{w} \in H^\perp \Rightarrow A\mathbf{w} \in H^\perp\,.$$

Proof. Let λ be the eigenvalue corresponding to the eigenvector $\mathbf{v}$. Suppose that $\mathbf{w} \in H^\perp$, that is, $\mathbf{w}\cdot\mathbf{v} = 0$. Then, by Lemma 7.27,

$$(A\mathbf{w})\cdot\mathbf{v} = \mathbf{w}\cdot(A\mathbf{v}) = \mathbf{w}\cdot(\lambda\mathbf{v}) = \lambda(\mathbf{w}\cdot\mathbf{v}) = 0\,,$$

so $A\mathbf{w} \in H^\perp$ as claimed. □

The following lemma contains the key result that will allow us to prove the Spectral Theorem. We already know from the discussion in the previous section that every matrix has a complex eigenvalue and a corresponding complex eigenvector. The following lemma shows that, rather unexpectedly, symmetric matrices always have at least one *real* eigenvalue with a corresponding real eigenvector:

Lemma 7.29. *Every symmetric matrix $A \in \mathbb{R}^{n\times n}$ has at least one real eigenvalue with corresponding real eigenvector $\mathbf{v} \in \mathbb{R}^n$.*

Proof. By Corollary 7.24 we know that A has at least one complex eigenvalue λ with corresponding eigenvector $\mathbf{z} \in \mathbb{C}^n$, that is

$$A\mathbf{z} = \lambda\mathbf{z}\,. \tag{7.12}$$

Write

$$\begin{aligned} &\lambda = a + ib \quad \text{where} \quad a, b \in \mathbb{R} \\ &\mathbf{z} = \mathbf{v} + i\mathbf{w} \quad \text{where} \quad \mathbf{v}, \mathbf{w} \in \mathbb{R}^n \end{aligned} \tag{7.13}$$

Thus, using (7.12) we have

$$A\mathbf{v} + iA\mathbf{w} = A(\mathbf{v} + i\mathbf{w}) = A\mathbf{z} = \lambda\mathbf{z} = (a + ib)(\mathbf{v} + i\mathbf{w}) = (a\mathbf{v} - b\mathbf{w}) + i(a\mathbf{w} + b\mathbf{v}),$$

which, by comparing real and imaginary parts, yields

$$\begin{aligned} A\mathbf{v} &= a\mathbf{v} - b\mathbf{w}, \\ A\mathbf{w} &= a\mathbf{w} + b\mathbf{v}. \end{aligned} \tag{7.14}$$

Now, by Lemma 7.27, we have

$$(a\mathbf{v} - b\mathbf{w})\cdot\mathbf{w} = (A\mathbf{v})\cdot\mathbf{w} = \mathbf{v}\cdot(A\mathbf{w}) = \mathbf{v}\cdot(a\mathbf{w} + b\mathbf{v}),$$

so

$$a(\mathbf{v}\cdot\mathbf{w}) - b\|\mathbf{w}\|^2 = a(\mathbf{v}\cdot\mathbf{w}) + b\|\mathbf{v}\|^2$$

and hence

$$b(\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2) = 0.$$

But $\mathbf{z} = \mathbf{v} + i\mathbf{w}$ is an eigenvector, so the vectors $\mathbf{v}$ and $\mathbf{w}$ cannot both be the zero vector. Therefore $\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 > 0$, and hence

$$b = 0.$$

Thus, using (7.13) and (7.14), we see that $\lambda = a \in \mathbb{R}$ and $A\mathbf{v} = a\mathbf{v} = \lambda\mathbf{v}$, that is, A has a real eigenvalue with corresponding real eigenvector. □

Proof of the Spectral Theorem. By induction on n. The theorem is trivial for $n = 1$. Suppose now that the theorem has been shown to be true for some k. Let $A \in \mathbb{R}^{(k+1)\times(k+1)}$ be symmetric. By Lemma 7.29 there is $\lambda \in \mathbb{R}$ and $\mathbf{v} \in \mathbb{R}^{k+1}$ such that

$$A\mathbf{v} = \lambda\mathbf{v}.$$

Since multiplying an eigenvector by a nonzero scalar does not change its status as an eigenvector, we may assume that $\mathbf{v}$ is normalised, that is, $\|\mathbf{v}\| = 1$. Next choose an orthonormal basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ for $\mathrm{Span}\,(\mathbf{v})^\perp$. Then $\{\mathbf{v}_1, \mathbf{w}_1, \ldots, \mathbf{w}_k\}$ is an orthonormal basis for $\mathbb{R}^{k+1}$ and it follows that the matrix $Q_0 \in \mathbb{R}^{(k+1)\times(k+1)}$ given by

$$Q_0 = \begin{pmatrix} \mathbf{v} & \mathbf{w}_1 & \cdots & \mathbf{w}_k \end{pmatrix}$$

is an orthogonal matrix. Now

$$Q_0^T A Q_0 = \begin{pmatrix} \mathbf{v}^T \\ \mathbf{w}_1^T \\ \vdots \\ \mathbf{w}_k \end{pmatrix} A \begin{pmatrix} \mathbf{v}^T & \mathbf{w}_1^T & \cdots & \mathbf{w}_k \end{pmatrix} = \begin{pmatrix} \mathbf{v}^T A\mathbf{v} & \mathbf{v}^T A\mathbf{w}_1 & \cdots & \mathbf{v}^T A\mathbf{w}_k \\ \mathbf{w}_1^T A\mathbf{v} & & & \\ \vdots & & B & \\ \mathbf{w}_k^T A\mathbf{v} & & & \end{pmatrix},$$

where $B = (b_{ij})$ is a real $k \times k$ matrix with

$$b_{ij} = \mathbf{w}_i^T A\mathbf{w}_j. \tag{7.15}$$

Moreover, since

$$\mathbf{v}^T A\mathbf{v} = \mathbf{v}^T(\lambda\mathbf{v}) = \lambda(\mathbf{v}^T\mathbf{v}) = \lambda,$$

and since, by Lemma 7.27 and Lemma 7.28,

$$\begin{aligned}\mathbf{v}^T A\mathbf{w}_j &= \mathbf{v}\cdot(A\mathbf{w}_j) = 0\\ \mathbf{w}_i^T A\mathbf{v} &= \mathbf{w}_i\cdot(A\mathbf{v}) = (A\mathbf{w}_i)\cdot\mathbf{v} = 0\,,\end{aligned}$$

it follows that

$$Q_0^T A Q_0 = \begin{pmatrix}\lambda & 0 & \cdots & 0\\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}.$$

Observe now that by Lemma 7.27 the $k\times k$ matrix B is symmetric, because

$$b_{ij} = \mathbf{w}_i^T A\mathbf{w}_j = \mathbf{w}_i\cdot(A\mathbf{w}_j) = (A\mathbf{w}_i)\cdot\mathbf{w}_j = \mathbf{w}_j\cdot(A\mathbf{w}_i) = \mathbf{w}_j^T A\mathbf{w}_i = b_{ji}\,.$$

Thus, by the inductive hypothesis, there is an orthogonal $k\times k$ matrix R such that

$$R^T B R = D\,,$$

with $D\in\mathbb{R}^{k\times k}$ diagonal. Define $Q_1\in\mathbb{R}^{(k+1)\times(k+1)}$ by

$$Q_1 = \begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & & & \\ \vdots & & R & \\ 0 & & & \end{pmatrix}.$$

A short calculation using the fact that $R^T R = I$ shows that $Q_1^T Q_1 = I$, so Q_1 is orthogonal. If we now let

$$Q = Q_0 Q_1\,,$$

then Q is orthogonal by Lemma 7.26, and

$$Q^T A Q = Q_1^T Q_0^T A Q_0 Q_1 = Q_1^T \begin{pmatrix}\lambda & 0 & \cdots & 0\\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix} Q_1 = \begin{pmatrix}\lambda & 0 & \cdots & 0\\ 0 & & & \\ \vdots & & R^T B R & \\ 0 & & & \end{pmatrix} = \begin{pmatrix}\lambda & 0 & \cdots & 0\\ 0 & & & \\ \vdots & & D & \\ 0 & & & \end{pmatrix},$$

so $Q^T AQ$ is diagonal. To summarise: we have just shown that if a $k\times k$ symmetric matrix can be diagonalised by an orthogonal matrix, then so can a $(k+1)\times(k+1)$ symmetric matrix. Combining this with the initial observation that, trivially, any 1×1 matrix can be diagonalised, the truth of the theorem now follows by the principle of induction. □

Let us record the following consequence of the Spectral Theorem, which is of independent interest:

Corollary 7.30. *The eigenvalues of a symmetric matrix A are real, and eigenvectors corresponding to distinct eigenvalues are orthogonal.*

Proof. Combine the Diagonalisation Theorem and the Spectral Theorem. □

Example 7.31. Consider the symmetric matrix

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

Find an orthogonal matrix Q that diagonalises A.

Solution. The characteristic polynomial of A is

$$\det(A - \lambda I) = -\lambda^3 + 3\lambda^2 + 9\lambda + 5 = (1+\lambda)^2(5-\lambda)\,,$$

so the eigenvalues of A are -1 and 5. Computing $N(A+I)$ in the usual way shows that $\{\mathbf{x}_1, \mathbf{x}_1\}$ is a basis for $N(A+I)$ where

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} .$$

Similarly, we find that the eigenspace $N(A-5I)$ corresponding to the eigenvalue 5 is 1-dimensional with basis

$$\mathbf{x}_3 = \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix} .$$

In order to construct the diagonalising orthogonal matrix for A it suffices to find orthonormal bases for each of the eigenspaces, since, by the previous corollary, eigenvectors corresponding to distinct eigenvalues are orthogonal.

In order to find an orthonormal basis for $N(A+I)$ we apply the Gram Schmidt process to the basis $\{\mathbf{x}_1, \mathbf{x}_2\}$ to produce the orthogonal set $\{\mathbf{v}_1, \mathbf{v}_2\}$:

$$\mathbf{v}_1 = \mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} ,$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} .$$

Now $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{x}_3\}$ is an orthogonal basis of $\mathbb{R}^3$ consisting of eigenvectors of A, so normalising them to produce

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \quad \mathbf{u}_2 = \frac{1}{\|\mathbf{v}_2\|}\mathbf{v}_2 = \begin{pmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}, \quad \mathbf{u}_3 = \frac{1}{\|\mathbf{x}_3\|}\mathbf{x}_3 = \begin{pmatrix} -1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix},$$

allows us to write down the orthogonal matrix

$$Q = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{3} & -1/\sqrt{6} \\ 0 & 1/\sqrt{3} & -2/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \end{pmatrix}$$

which diagonalises A, that is,

$$Q^T A Q = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{pmatrix} .$$

□

Index