UNIVERSITY OF STRATHCLYDE
DEPARTMENT OF MATHEMATICS AND STATISTICS

Geometry and Linear Algebra
Chapter 1: Introduction

Geometry is the branch of mathematics concerned with the relationships between and properties of objects in space. The relationships and properties include symmetries, distances and angles and the objects considered are (conventionally) points, curves, surfaces or solids.

The history of geometry is almost as long as human history itself, with evidence of its development in the ancient civilisations of Egypt, India and China, amongst others. Initially, geometry was one of the most applied of mathematical subjects, with its obvious connections with people's place in the world around them through measurement of relative distance, area and volume. But it has also long been used by pure mathematicians as a vehicle for abstraction: the point where you meet a friend becomes an infinitesimally small location in mathematical space.

Around 2500 years ago, the great Greek mathematician Euclid popularised an axiomatic approach to geometry, establishing complex relationships in space using a handful of self-evident properties (such as "a straight line can be drawn between any two points"). His books, entitled "The Elements" are full of long wordy proofs and complex diagrams (see Figure 1.1) but he was working without most of the tools we now take for granted.

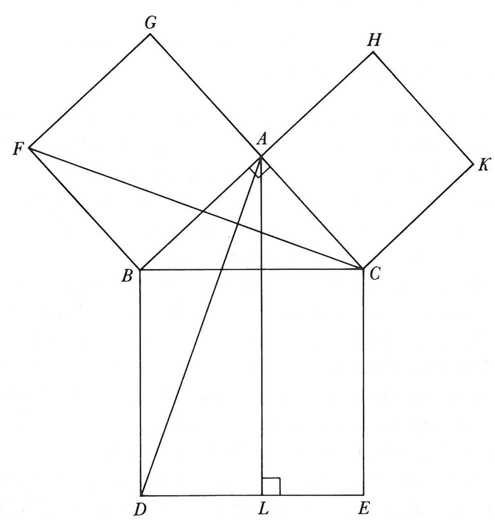


Figure 1.1: Pythagoras' theorem by Euclid.

By representing geometrical concepts **algebraically** and using **coordinate systems** we can manipulate objects in space indirectly, and this often leads to great simplification as well as giving insight into deeper relationships which a direct approach can obscure, and to give a handle on higher dimensional objects which the mind cannot picture.

A particular focus of this course is **linear transformations** of space. Such transformations can be represented algebraically using **matrices** which can be manipulated through **matrix algebra**, a major area of linear algebra. Matrices (and linear transformations) have applications that range widely through almost all areas of mathematics. For example, differentiation and integration can be viewed as linear transformations. Hopefully, you'll be able to apply the ideas you learn in this course throughout your future mathematical studies.

Despite its long history, geometry is still an evolving subject. A key insight was the realisation that Euclid's "self-evident properties" need not always hold. What happens if we let parallel lines meet? In the Euclidean plane this can only happen at infinity, but in other geometries it doesn't create a problem at all: think of lines of latitude on the globe. Classical geometry has spawned many different branches of higher mathematics. Examples include differential geometry (with applications ranging from general relativity to computer animation), noncommutative geometry (a key mathematical tool in quantum physics) and topology (a celebrated problem in this field recently earned its solver $1,000,000). You may encounter some of these areas in your future studies, but in this course we will restrict ourselves to more familiar behaviour in two and three dimensional Euclidean space. To start off, we introduce some of the tools we'll need to understand these spaces.

1.1 Cartesian Coordinates

As a geometer, one may want to describe relations between the points and sides of an object such as the triangle in Figure 1.2. The geometrical techniques espoused by Euclid for answering geometrical questions, although exhibiting a rigour previously unseen, can be cumbersome; and it often appears that each proof needs a new and ingenious approach.

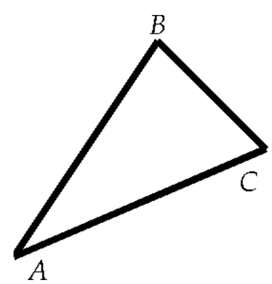


Figure 1.2: A triangle.

Coordinate geometry was introduced by the French philosopher René Descartes (most famous for his maxim "cogito ergo sum"). His motivation was to allow geometry to be approached methodically by resolving geometrical questions algebraically. His system of Cartesian coordinates (Descartes was also known by the Latin name "Cartesius") assigns a label to every point in the Euclidean plane describing the positions in space of points relative to an origin O. In two dimensions this can be done with a pair of coordinates representing horizontal (the x direction) and vertical distance (the y direction). Notice that the resulting axes are at right angles. Given the coordinates of the points A, B and C one can work out the distance and angles between them with algebraic formulae.

For example, using Figure 1.3 we see that A has coordinates $(0, 3)$ and B has coordinates $(4, 9)$, so the

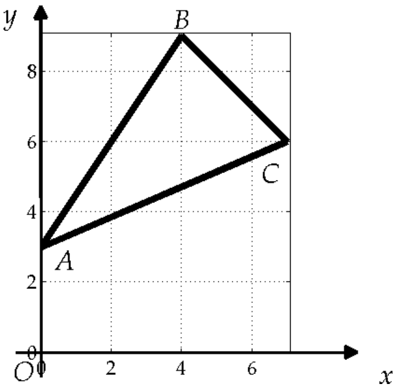


Figure 1.3: Cartesian coordinates.

distance between the two points (or side length) is given by

$$d_{AB} = \sqrt{(4-0)^2 + (9-3)^2} = \sqrt{52}.$$

Here we have applied the **distance formula**. If P has coordinates (x, y) and Q has coordinates (ξ, η) then the Euclidean distance between the two points is given by the formula

$$d_{PQ} = \sqrt{(x-\xi)^2 + (y-\eta)^2}. \tag{1.1.1}$$

A geometrical fact given by a purely algebraic formula. Angles, too, can be expressed algebraically from the coordinates, a theme we'll return to (you may remember a formula for angles involving the scalar product).

Coordinates allow us to describe all sorts of curves algebraically. An equation such as $y = f(x)$ can be interpreted as a relationship that describes a particular set of points in the plane. Other curves may be described by **implicit** expressions such as $g(x, y) = 0$ or **parametric** expressions, such as $x = p(t), y = q(t)$.

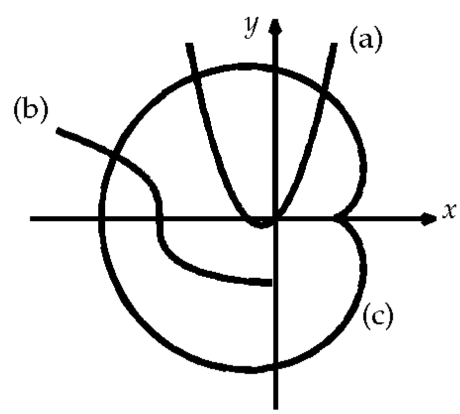
Example 1.1.1

- Draw the curves described by the following expressions.

  (a) $y = 2x^2 + x$, $x \in [-3/2, 1]$.

  (b) $x^2 - 3y^3 = 4$, $x \in [-4, 0]$.

  (c) $x = 2\cos t - \cos 2t$, $y = 2\sin t - \sin 2t$, $0 \le t \le 2\pi$.



1.1.1 Moving The Origin

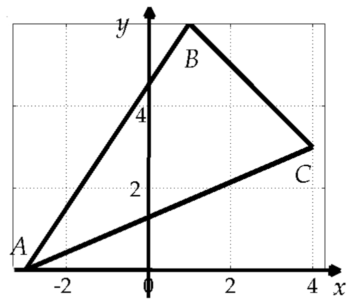


Figure 1.4: Changing coordinates.

The coordinates we use in Figure 1.3 are, to a certain extent, arbitrary. In Figure 1.4 we have moved the origin relative to the triangle (up and to the right). This hasn't changed the triangle: its sides are the same length, the angles are unchanged and geometric results to do with lengths and angles should remain unchanged, too. But look now at Figure 1.5. Here we have kept the origin the same, but changed the spacing between points. The shape has not changed: angles are preserved and the triangle looks identical but if we use (1.1.1) we'll find the side lengths have changed.

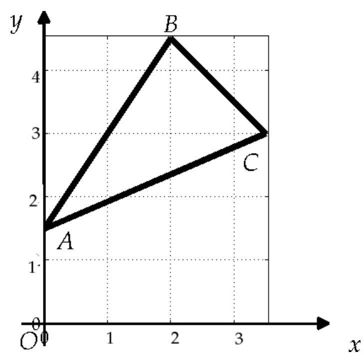


Figure 1.5: Stretching coordinates.

Later on, we'll look more closely at properties that are preserved when we change axes. For now, just consider what Figure 1.6 illustrates.

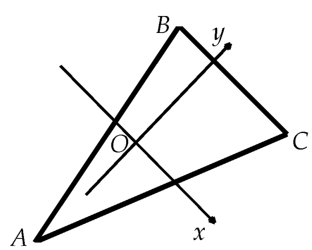


Figure 1.6: Rotating coordinates.

1.1.2 Pythagoras' Theorem

The distance formula (1.1.1) is a consequence of Pythagoras's theorem. Figure 1.1 illustrates Euclid's proof. The fact that the area of the two smaller squares equals that of the larger one follows from a sequence of deductions. For example, we start by noting that the triangles CBF and ABD have the same area (why? Compare the lengths BF and AB, and BC and BD. Also note that the angles CBF and ABD are both made up of angle CBA plus a right angle).

However, algebraic proofs can simplify the matter. Look at Figure 1.7.

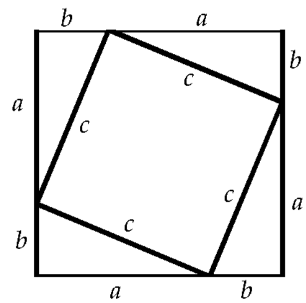


Figure 1.7: Pythagoras revisited.

Consider a right angled triangle with side lengths a, b and c. The big square has area $(a+b)^2$ and is made of the interior square (How do we know this is a square?) and four right angled triangles. Thus

$$a^2 + 2ab + b^2 = c^2 + 4(ab/2),$$

and $a^2 + b^2 = c^2$, as desired. While the coordinates are implicit in this proof, it reinforces the notion that geometrical understanding is enhanced by algebra.

1.1.3 Polar Coordinates

In two dimensions, we need to assign two coordinates to each point in order to describe them unambiguously. There are many options for assigning these values, and we are not limited to the Cartesian approach of defining coordinates according to horizontal and vertical distance. One particularly useful alternative is to use polar coordinates.

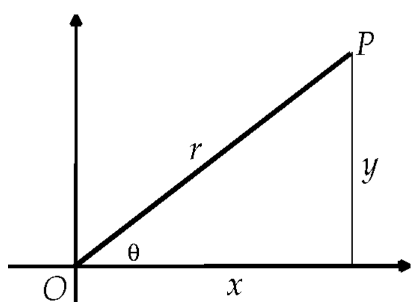


Figure 1.8: Polar coordinates.

The point P is specified uniquely by fixing the values of x and y but it can also be specified by giving values to $r \geq 0$ (the distance OP) and θ (the angle between OP and the positive x-axis). These values are P's **polar coordinates**.

Using basic trigonometry we can easily convert from one set of coordinates to the other.

$$x = r\cos\theta, \qquad y = r\sin\theta,$$

converts polar to Cartesian.

$$r = \sqrt{x^2 + y^2}, \qquad \tan\theta = \frac{y}{x},$$

converts Cartesian to polar if we solve for θ (taking careful note of the quadrant our point lies in).

Polar coordinates are particularly useful for describing problems with radial symmetry.

Examples 1.1.2

Exercises: Cartesian Coordinates

1. Draw the curves described by the following expressions.

   (a) $y = 3\sin x - 2x^2 + 3x$, $x \in [0, 2]$ (remember to use radians!).

   (b) $x^2 - y^2 = 1$, $x \in [-2, 2]$.

   (c) $x = \cos 3t$, $y = \sin 2t$, $0 \leq t \leq \pi$.

2. The triangle ABC has vertices whose coordinates are $A(0,0)$, $B(3,0)$ and $C(1,1)$. DEF is the triangle with vertices $D(1,0)$, $E(4,0)$ and $F(2,-1)$. Sketch the two triangles and confirm that they are **congruent**: they are the same shape and size (just oriented differently). Confirm that ABC can be transformed into DEF by translating it one unit along the x-axis and then reflecting through the x-axis.

   Draw the following pairs of congruent triangles and describe a set of transformations to take one into the other: you'll need a combination of translations, rotations and/or reflections.

   (a) $A(0,0)$, $B(2,0)$, $C(2,1)$ and $D(0,2)$, $E(-1,2)$, $F(0,0)$.

   (b) $A(6,2)$, $B(-1,3)$, $C(-3,5)$ and $D(-1,3)$, $E(8,0)$, $F(1,1)$.

   (c) $A(2,1)$, $B(5,2)$, $C(3,5)$ and $D(1,2)$, $E(2,-1)$, $F(5,1)$.

   (d) $A(0,0)$, $B(2,1)$, $C(1,3)$ and $D(-1,-1)$, $E(0,1)$, $F(2,0)$.

   (e) $A(0,0)$, $B(1,0)$, $C(1/2, \sqrt{3}/2)$ and $D(0,0)$, $E(-1/2, \sqrt{3}/2)$, $F(1/2, \sqrt{3}/2)$.

3. Find the polar form of the following points given in Cartesian coordinates.

   (a) $(2, -\sqrt{12})$ (b) $(3,4)$ (c) $(-12,0)$ (d) $(-128,-128)$ (e) $(6,6)$
   (f) $(0,4)$ (g) $(0,-2)$ (h) $(-3, \sqrt{3})$.

4. Convert the polar coordinates of the following points to Cartesian form and sketch them.

   (a) $(r,\theta) = (3, \frac{15}{4}\pi)$ (b) $(r,\theta) = (6, \frac{11}{6}\pi)$ (c) $(r,\theta) = (2, 7\pi)$.

5. The shape of the infinity sign, ∞, is known as a lemniscate. The lemniscate of Bernoulli is described in Cartesian form by the expression $(x^2 + y^2)^2 = x^2 - y^2$.

   (a) By converting Cartesian coordinates to polar coordinates, find the polar form of the lemniscate.

   (b) The parametric form of the lemniscate is

   $$x = \frac{\cos t}{1 + \sin^2 t}, \quad y = \frac{\sin 2t}{2 + 2\sin^2 t}, \quad 0 \leq t \leq 2\pi.$$

   Sketch the curve.

   (c) Substitute the parametric form of the curve into the polar and Cartesian expressions to confirm their validity.

1.2 Vectors

A **vector** is a mathematical quantity that has both magnitude and direction. Quantities which have just a magnitude (such as numbers) are called **scalars**. Just as is the case for numbers, vectors have many roles in mathematics. Geometrically, we will use vectors to represent directions and positions. But we will also manipulate them algebraically and the **components** of vectors can represent almost any quantity. When working with vectors it is vital to be clear what they represent: positions, directions or general quantities. In its most general form, a vector can be a very abstract object. Our starting point will be the following algebraic definition.

Definition 1.2.1 *A vector* $\mathbf{v}$ *of dimension* n *with components* $v_1, v_2, \ldots, v_n$ *is written*

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix},$$

where $v_i \in \mathbb{R}$, $i = 1, 2, \ldots, n$, $-\mathbf{v}$ *is the vector whose i-th component is* $-v_i$ *and for any scalar* $k \in \mathbb{R}$, $k\mathbf{v}$ *is the vector whose i-th component is* kv_i.

If $\mathbf{v}$ *and* $\mathbf{w}$ *are vectors* **of the same dimension** *then* $\mathbf{v} + \mathbf{w}$ *is the vector whose i-th component is* $v_i + w_i$ *and* $\mathbf{v} = \mathbf{w}$ *if and only if* $v_i = w_i$ *for* $i = 1, 2, \ldots, n$.

The vector $\mathbf{0}$ *(of dimension n) is the vector all of whose components are zero.*

Most of the results that hold for our vectors can be generalised to more abstract definitions of vectors, where individual components can be much more elaborate—polynomials, for example—but geometrical results are more easily understood with our definition. Almost all the vectors we will use will be 2 or 3 dimensional. We can represent a vector $\mathbf{v}$ by a line joining the origin to the point with coordinates $(v_1, v_2, \ldots, v_n)$ with an arrow to denote direction. This connection between the geometry and algebra of vectors will inform us throughout the course.

1.2.1 Position Vectors

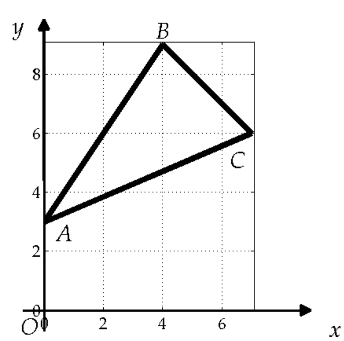


Figure 1.9: Return of the triangle.

Here's a picture we've seen before of a triangle whose vertices A, B and C have coordinates $(0, 3)$, $(4, 9)$ and $(7, 6)$, respectively. Let us describe this picture with vectors. The vectors

$$\begin{bmatrix} 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 9 \end{bmatrix}, \begin{bmatrix} 7 \\ 6 \end{bmatrix}$$

are the **position vectors** of the vertices. We use the notation $\overrightarrow{OP}$ to represent the position vector of a point P (sometimes we will use $\mathbf{p}$ instead), thus

$$\overrightarrow{OA} = \mathbf{a} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \ \overrightarrow{OB} = \mathbf{b} = \begin{bmatrix} 4 \\ 9 \end{bmatrix}, \ \overrightarrow{OC} = \mathbf{c} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}.$$

Now consider the vector

$$\mathbf{d} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}.$$

This is the position vector of the point D with coordinates $(4, 6)$. But it is also the direction that takes us from A to B. To denote this fact we can write $\mathbf{d} = \overrightarrow{AB}$. We use equality to denote that the vectors have identical length and direction even though they are located in different positions.

From Figure 1.9 we can ascertain the following

$$\overrightarrow{AB} = \mathbf{b} - \mathbf{a} = -\overrightarrow{BA}, \ \ \overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}.$$

It should be evident that these results are true for any three points A, B and C.

We can rewrite Definition 1.2.1 from a geometrical view point as follows.

Definition 1.2.2 *A vector* $\mathbf{v}$ *of dimension* n *with components* $v_1, v_2, \ldots, v_n$ *is the straight line path* $\overrightarrow{OV}$ *from the origin to the point* V *with coordinates* $(v_1, v_2, \ldots, v_n)$. *For any scalar* $k > 0$, $k\mathbf{v}$ *is the vector in the same direction as* $\mathbf{v}$ *whose length is* k *times that of* $\mathbf{v}$ *while* $-k\mathbf{v}$ *has the same length as* $k\mathbf{v}$ *but points in the opposite direction, hence* $-\mathbf{v} = \overrightarrow{VO}$.

If $\mathbf{v}$ *and* $\mathbf{w}$ *are vectors of the same dimension then* $\mathbf{v} + \mathbf{w}$ *is the position vector of the point whose i-th coordinate is* $v_i + w_i$, *forming the fourth corner of a parallelogram whose other corners are the origin,* $(v_1, v_2, \ldots, v_n)$ *and* $(w_1, w_2, \ldots, w_n)$. $\mathbf{v} = \mathbf{w}$ *if and only if* $\mathbf{v}$ *and* $\mathbf{w}$ *are in identical directions and of equal length.*

The vector $\mathbf{0}$ *(of dimension* n*) is a vector of zero length and indeterminate direction.*

Examples 1.2.1

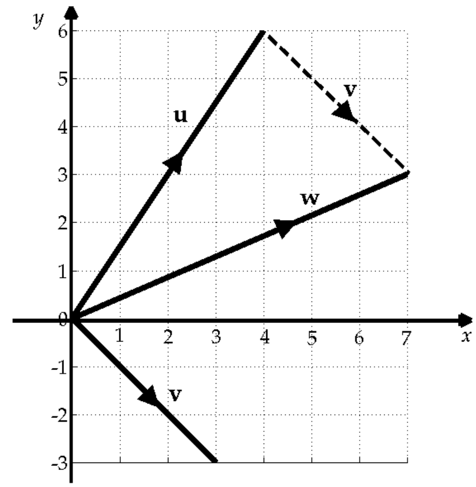
(i) From Figure 1.9 we see that

$$\overrightarrow{AB} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}, \ \overrightarrow{BC} = \begin{bmatrix} 3 \\ -3 \end{bmatrix}, \ \overrightarrow{AC} = \begin{bmatrix} 7 \\ 3 \end{bmatrix}.$$

As is to be expected,

$$\overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CA} = \begin{bmatrix} 4 \\ 6 \end{bmatrix} + \begin{bmatrix} 3 \\ -3 \end{bmatrix} - \begin{bmatrix} 7 \\ 3 \end{bmatrix} = \begin{bmatrix} 4+3-7 \\ 6-3-3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Notice that the same equalities hold true (with the same numbers) for Figure 1.4. We can also illustrate this with position vectors.

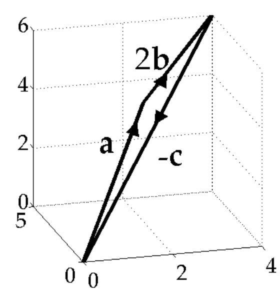


Moving the position vector $\mathbf{v}$ as we have done in our diagram makes it easy to see that $\mathbf{u} + \mathbf{v} - \mathbf{w} = \mathbf{0}$.

(ii) Everything we have seen so far can be extended to vectors in higher dimension. For example, if

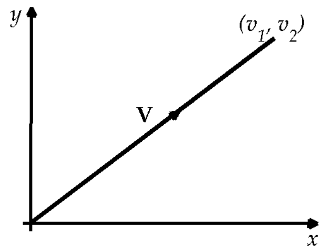
$$\mathbf{a} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix},$$

then $\mathbf{a} + 2\mathbf{b} - \mathbf{c} = \mathbf{0}$.



Note that expressions such as $\mathbf{a} + 2\mathbf{b} - \mathbf{c}$, $\overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CA}$ and $k\mathbf{v}$ are known as **linear combinations** of vectors.

1.2.2 Length of A Vector



An obvious definition of the length of the vector **v** is the straight line distance between its end points. For example, in two dimensions Pythagoras' theorem tells us that $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ has length $\sqrt{v_1^2 + v_2^2}$. This can be generalised to any dimension.

Definition 1.2.3 *If*

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix},$$

then the **length** *of* **v**, *written* $\|\mathbf{v}\|$, *is*

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

The length is also known as the **norm** *or* **Euclidean norm**.

Example 1.2.2

- The norms of the vectors in Example 1.2.1(i) are

$$\|\overrightarrow{AB}\| = \sqrt{52}, \quad \|\overrightarrow{BC}\| = \sqrt{18} = 3\sqrt{2}, \quad \|\overrightarrow{AC}\| = \sqrt{58}.$$

  In other words, the distances between A and B, B and C, and C and A are $\sqrt{52}$, $3\sqrt{2}$ and $\sqrt{58}$, respectively.

- The lengths of the vectors in Example 1.2.1(ii) are

$$\|\mathbf{a}\| = \sqrt{29}, \quad \|\mathbf{b}\| = \sqrt{3}, \quad \|\mathbf{c}\| = \sqrt{77}.$$

- Suppose we have a triangle ABC where $\mathbf{u} = \overrightarrow{AB}$, $\mathbf{v} = \overrightarrow{BC}$ and $\mathbf{w} = \overrightarrow{AC}$. For any triangle, the **triangle inequality** states that the length of one side must be less than or equal to the sum of the other two sides. Comparing AC to AB and BC we can write this in vector notation as

$$\|\mathbf{w}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|,$$

  and since $\mathbf{w} = \mathbf{u} + \mathbf{v}$, we can write the triangle inequality as

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

  This is a tremendously useful result.

From the definition of the norm we can see that $\|\mathbf{v}\| \geq 0$ for all vectors with equality if and only if $\mathbf{v} = \mathbf{0}$.

If the norm of a vector equals one then it is called a **unit vector**. If $\mathbf{v}$ is any nonzero vector then a unit vector in the direction of $\mathbf{v}$ is given by $\hat{\mathbf{v}} = \mathbf{v}/\|\mathbf{v}\|$. Note that there are two unit vectors in any direction: if $\hat{\mathbf{v}}$ is a unit vector then so is $-\hat{\mathbf{v}}$.

In two dimensions, two particularly important unit vectors are

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{1.2.1}$$

Their three dimensional counterparts are

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{1.2.2}$$

and so on for higher dimensions. Notice the meaning of $\mathbf{e}_i$ is dependent on the dimension, but the context should make this meaning clear. A common alternative notation in three dimensions is $\mathbf{i} = \mathbf{e}_1$, $\mathbf{j} = \mathbf{e}_2$ and $\mathbf{k} = \mathbf{e}_3$.

Example 1.2.3

- Unit vectors in the direction of those in Examples 1.2.1(i) are

$$\hat{\mathbf{u}}_1 = \begin{bmatrix} \frac{4}{\sqrt{52}} \\ \frac{6}{\sqrt{52}} \end{bmatrix}, \quad \hat{\mathbf{u}}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad \hat{\mathbf{u}}_3 = \begin{bmatrix} \frac{7}{\sqrt{58}} \\ \frac{3}{\sqrt{58}} \end{bmatrix},$$

  and for those in Examples 1.2.1(ii),

$$\hat{\mathbf{a}} = \frac{1}{\sqrt{29}} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}, \quad \hat{\mathbf{b}} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \hat{\mathbf{c}} = \frac{1}{\sqrt{77}} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}.$$

1.2.3 Scalar Product

The norm allows us to find relationships between the lengths of vectors. Another useful relationship is the angle between vectors. A straightforward way of finding angles is to use the **scalar product**.

Look at Figure 1.10 and suppose that the angles between the x-axis and $\mathbf{u}$ and $\mathbf{v}$ are ϕ and ψ, respectively. Then, since we can use the norm to measure the vector length,

$$u_1 = \|\mathbf{u}\| \cos\phi, \quad u_2 = \|\mathbf{u}\| \sin\phi, v_1 = \|\mathbf{v}\| \cos\psi, \quad v_2 = \|\mathbf{v}\| \sin\psi.$$

So,

$$\begin{aligned} \cos\theta &= \cos(\phi - \psi) = \cos\phi\cos\psi + \sin\phi\sin\psi \\ &= \frac{u_1}{\|\mathbf{u}\|} \times \frac{v_1}{\|\mathbf{v}\|} + \frac{u_2}{\|\mathbf{u}\|} \times \frac{v_2}{\|\mathbf{v}\|} \\ &= \frac{u_1 v_1 + u_2 v_2}{\|\mathbf{u}\|\|\mathbf{v}\|}. \end{aligned}$$

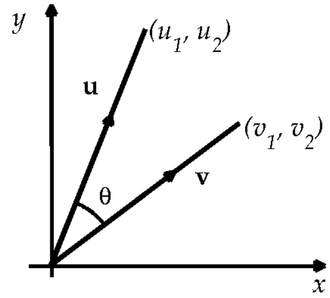


Figure 1.10: Angle between vectors.

Definition 1.2.4 *If*

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix},$$

then the scalar product of $\mathbf{u}$ *and* $\mathbf{v}$*, written* $\mathbf{u} \cdot \mathbf{v}$ *is defined to be*

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n. \tag{1.2.3}$$

Notice that the scalar product can only be taken between vectors of the same dimension. The scalar product is also know as the **dot product** or the **inner product**. As the name suggests, the result of this operation between two vectors is a scalar. Alternative notations for the scalar product include $\mathbf{u}^T\mathbf{v}$ (we'll use this later) and $\langle \mathbf{u}, \mathbf{v} \rangle$ (which you'll encounter in future classes).

The relationship between angles and scalar products holds in higher dimensions, too. That is, the angle θ between two vectors $\mathbf{u}$ and $\mathbf{v}$ emerging from a common point satisfies

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}.$$

By rearranging this expression we get the geometric definition of the scalar product

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\| \cos\theta. \tag{1.2.4}$$

A useful consequence of this relationship is that if $\mathbf{u} \cdot \mathbf{v} = 0$ then $\mathbf{u} = \mathbf{0}$, $\mathbf{v} = \mathbf{0}$, or (more usually) the vectors are at right angles to each other.

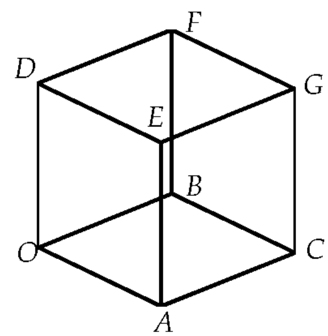
Notice that there is more than one way for the angle between two vectors to be measured. For example, in Figure 1.10 the angle between $\mathbf{u}$ and $\mathbf{v}$ could be written θ, $-\theta$, $2\pi - \theta$, or $\theta - 2\pi$. However, as the cosine of all of these angles is the same it makes no difference to the definition of the scalar product. The angle can always be written uniquely by choosing $\theta \in [0, \pi]$ (or $[0, 180]$ if using degrees). A positive scalar product indicates an acute angle between vectors and a negative value means the angle is obtuse.

Examples 1.2.4

Examples 1.2.5

(i) In the unit cube $OABCDEFG$, find the angle between OA and the diagonal OG.
Using Cartesian coordinates we can write

$$\overrightarrow{OA} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \overrightarrow{OG} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$



hence

$$\cos \angle GOA = \frac{\overrightarrow{OA} \cdot \overrightarrow{OG}}{\|\overrightarrow{OA}\| \|\overrightarrow{OG}\|} = \frac{1}{\sqrt{3}},$$

and $\angle GOA \approx 0.96$ radians.

(ii) A father and his daughter are playing on a see-saw. The father weighs 80kg and his daughter weighs 25kg. The distance from the end of the see-saw to the middle is 4m. If the daughter sits at one end, how far from the centre should the father sit to balance the see-saw?

What about a father who weighs $w_1$kg and a daughter who weighs $w_2$kg on a see-saw which is $2d_2$m from end-to-end?

What does this have to do with scalar products?

To balance, the moment generated by the man and girl must be equal. So $80d_1 = 25 \times 4$, where d_1 is the man's position. So $d_1 = 1.25$m. To be more precise, we should write $d_1 = -1.25$m as the man should be the opposite side of the centre from the girl.

In the general problem, $d_1 = -w_2 d_2 / w_1$, or $w_1 d_1 + w_2 d_2 = 0$. Letting

$$\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

be the vectors of distances and weights (or masses), the moment balancing equation becomes

$$\mathbf{d} \cdot \mathbf{w} = 0.$$

Exercises: Vectors

1. Draw the following vectors to scale on the same diagram.

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \end{bmatrix}, \begin{bmatrix} -3 \\ 2 \end{bmatrix}.$$

For questions 2–7, let $\mathbf{a} = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$ and $\mathbf{d} = \begin{bmatrix} -1 \\ -3 \\ -2 \end{bmatrix}$.

2. Evaluate the following expressions.

   (a) $\mathbf{a} + \mathbf{b} - 3\mathbf{c}$ (b) $2\mathbf{d} - \mathbf{b}$ (c) $\mathbf{c} \cdot \mathbf{d}$ (d) $\mathbf{c} \cdot (2\mathbf{d} - 3\mathbf{b})$ (e) $2\mathbf{a} - \mathbf{b} + \mathbf{c} + \mathbf{d}$.

3. (a) Find scalars a, b and c so that the linear combination $a\mathbf{a} + b\mathbf{b} + c\mathbf{c} = \mathbf{e}_1$, where $\mathbf{e}_1$ is defined in (1.2.2).

   (b) Use 2(e) to find a linear combination of $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ that equals $\mathbf{e}_1$.

4. Let $\mathbf{f} = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$. Find a linear combination of $\mathbf{a}$ and $\mathbf{b}$ that equals $\mathbf{f}$.

   Can you find a linear combination of $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{f}$ that equals $\mathbf{e}_1$?

5. Find the smallest angle between $\mathbf{a}$ and, respectively, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$.

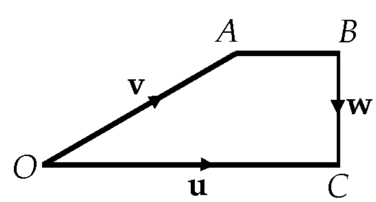
6. Find the norms of the vectors in 2(a)(b)(e).

7. Find unit vectors in the directions of $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$.

8. Vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are represented on the diagram shown where the angle $AOC = \pi/6$ radians.

   (a) Given that $\|\mathbf{u}\| = 4$ and $\|\mathbf{v}\| = 3$, evaluate $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w})$ and $\mathbf{w} \cdot (\mathbf{u} - \mathbf{v})$.

   (b) Find $\|\mathbf{v} + \mathbf{w}\|$ and $\|\mathbf{u} - \mathbf{v}\|$.



9. Indicate whether the following statements are true or false. Justify your answers.

   (a) If $\mathbf{x} + \mathbf{y} = \mathbf{x} + \mathbf{z}$ then $\mathbf{y} = \mathbf{x}$.

   (b) If $\mathbf{x} + \mathbf{y} = \mathbf{0}$ then $c\mathbf{x} + d\mathbf{y} = \mathbf{0}$ for all c and d.

   (c) Parallel vectors of the same length are equal.

   (d) If $a\mathbf{x} = \mathbf{0}$ then either $a = 0$ or $\mathbf{x} = 0$.

   (e) If $c\mathbf{x} + d\mathbf{y} = \mathbf{0}$ then $\mathbf{x}$ and $\mathbf{y}$ are parallel vectors.

10. Suppose you are a stockbroker who deals in three different shares. Suppose the share prices are (in pounds) p_1, p_2 and p_3 and the quantities you buy and sell are q_1, q_2 and q_3 (positive if you buy and negative if you sell). One day your income from share deals is £1 million. Express this profitable day using scalar products.

11. Show that for any vector $\mathbf{u}$ and any scalar a, $\|a\mathbf{u}\| = |a|\|\mathbf{u}\|$.

12. Show that if $k > 0$ and $\mathbf{y} = k\mathbf{x}$ then $\|\mathbf{x} + \mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$.

    Draw a picture to explain why the triangle inequality is an equality in this case.

13. Show that $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathbf{u} \cdot \mathbf{v}$.

    Use (1.2.4) to show that the triangle inequality is an equality if and only if $\mathbf{u}$ and $\mathbf{v}$ point in the same direction.

14. Use the triangle inequality to derive the reverse triangle inequality $\|\mathbf{u} - \mathbf{v}\| \geq \Big| \|\mathbf{u}\| - \|\mathbf{v}\| \Big|$.

15. Let $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Sketch a picture showing all vectors $\mathbf{w}$ such that $\mathbf{v} \cdot \mathbf{w} = 5$. Which is the shortest $\mathbf{w}$?

16. (a) Let $\mathbf{a} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Calculate $\mathbf{p} = \dfrac{\mathbf{a} \cdot \mathbf{x}}{\|\mathbf{x}\|^2}\mathbf{x}$ and $\mathbf{q} = \mathbf{a} - \mathbf{p}$.

    Sketch $\mathbf{a}$, $\mathbf{x}$, $\mathbf{p}$ and $\mathbf{q}$ on a diagram. Comment on what you see.

    (b) Given any two nonzero vectors $\mathbf{a}$ and $\mathbf{x}$, the vector $\mathbf{p} = \dfrac{\mathbf{a} \cdot \mathbf{x}}{\|\mathbf{x}\|^2}\mathbf{x}$ is known as the **projection** of $\mathbf{a}$ onto $\mathbf{x}$. Show that $\mathbf{a} - \mathbf{p}$ and $\mathbf{x}$ meet at right angles.

1.3 Matrices

A matrix is a rectangular array of numbers. If the matrix has m rows and n columns it is called an $m \times n$, or m-by-n, matrix. These bland statements do not do justice to the fundamental role matrices play in pure and applied mathematics by, in effect, generalising the idea of a number to higher dimensions. This is a rather abstract notion, but we will see its impact in plenty of practical situations. In order for our multi-dimensional "numbers" to work, we need to define the basic arithmetical operations. Initially we'll define addition, subtraction and, after a couple of additional definitions, multiplication. Division comes later.

We will denote matrices by capital letters. If A is an $m \times n$ matrix the number in the ith row and jth column is referred to as its (i, j)-th entry (or component) and denoted by a_{ij} or $A(i, j)$.

Definition 1.3.1 *Suppose that A and B are both $m \times n$ matrices.*

$C = A + B$ is an $m \times n$ matrix whose (i, j)-th entry is $c_{ij} = a_{ij} + b_{ij}$. That is, we add matrices together **componentwise** *and we can only add matrices together of the same dimension.*

The **transpose** *of the $m \times n$ matrix A, written A^T is the $n \times m$ matrix whose (i, j)-th entry is a_{ji}.*

For any scalar $k \in \mathbb{R}$, kA is the matrix whose (i, j)-th entry is ka_{ij}. Multiplication of a matrix by a scalar is componentwise. $D = A - B = A + (-B)$ is the matrix whose (i, j)-th entry is $d_{ij} = a_{ij} - b_{ij}$.

The $m \times n$ zero matrix, written O, is the matrix all of whose entries are zero. $A + B = A$ if and only if $B = O$. $A - B = O$ if and only if $A = B$.

Examples 1.3.1

(i) $\begin{bmatrix} 4 & 6 & 3 \\ 2 & 5 & 3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 6 & 5 \\ 3 & 7 & 3 \end{bmatrix}$.

(ii) We cannot add $\begin{bmatrix} 2 & 0 & 3 \\ 0 & 2 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ together because their dimensions do not agree.

(iii) Let $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$. Then $3A = \begin{bmatrix} 3 & 6 \\ 0 & 3 \end{bmatrix}$ and $A^T = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$.

Example 1.3.2

- A vector with n components can be thought of as an $n \times 1$ matrix. Vector algebra and matrix algebra do not contradict each other: Definition 1.3.1 generalises Definition 1.2.1.
- Suppose $\mathbf{v} = \begin{bmatrix} 2 \\ 0 \\ 5 \end{bmatrix}$. Then $\mathbf{v}^T = \begin{bmatrix} 2 & 0 & 5 \end{bmatrix}$. We call $\mathbf{v}$ a **column vector** and $\mathbf{v}^T$ a **row vector**.

- Let $A = \begin{bmatrix} 1 & 2 \\ 6 & 7 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix}$. Then

$$(A^T)^T = \begin{bmatrix} 1 & 6 \\ 2 & 7 \end{bmatrix}^T = A, \quad (A+B)^T = \begin{bmatrix} 2 & 4 \\ 5 & 9 \end{bmatrix}^T = \begin{bmatrix} 2 & 5 \\ 4 & 9 \end{bmatrix} = A^T + B^T.$$

  The results $(A^T)^T = A$ and $(A+B)^T = A^T + B^T$ are true for all matrices.

- Let $A = \begin{bmatrix} 0 & 3 & 4 \\ 3 & 7 & 2 \\ 4 & 2 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 5 & -2 \\ -5 & 0 & 1 \\ 2 & -1 & 0 \end{bmatrix}$. Then $A = A^T$ and $B = -B^T$.

- Let $A = \begin{bmatrix} 1 & 7 & 5 \\ 3 & 6 & -1 \\ -3 & 1 & -2 \end{bmatrix}$. Let $B = (A + A^T)/2$ and $C = (A - A^T)/2$. Then,

$$\begin{aligned}
B + C &= A/2 + A^T/2 + A/2 - A^T/2 = A, \\
B &= \frac{1}{2}\begin{bmatrix} 1 & 7 & 5 \\ 3 & 6 & -1 \\ -3 & 1 & -2 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 & 3 & -3 \\ 7 & 6 & 1 \\ 5 & -1 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 5 & 1 \\ 5 & 6 & 0 \\ 1 & 0 & -2 \end{bmatrix}, \\
C &= \frac{1}{2}\begin{bmatrix} 1 & 7 & 5 \\ 3 & 6 & -1 \\ -3 & 1 & -2 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 & 3 & -3 \\ 7 & 6 & 1 \\ 5 & -1 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 4 \\ -2 & 0 & -1 \\ -4 & 1 & 0 \end{bmatrix}.
\end{aligned}$$

  Notice that $B = B^T$, $C = -C^T$.

Here is some useful terminology.

- A matrix is **square** if it has as many rows as columns (and is then an $n \times n$ matrix).
- If $A = A^T$ then we call A a **symmetric matrix**. If $A = -A^T$ then we call A a **skew-symmetric matrix**. Only square matrices can be symmetric (or skew-symmetric). The last example generalises: any square matrix can be written as the sum of a symmetric and a skew-symmetric matrix.
- The **diagonal** of a square matrix A is the set of entries a_{ii}. A **diagonal matrix** is a matrix whose only nonzero entries lie along the diagonal.
- The rows and columns of a matrix are vectors. We use the notation $\mathbf{a}_j$ to denote the j-th column of a matrix and $\mathbf{a}_i^T$ to denote the i-th row. Beware: there is scope for ambiguity here, but the context should be clear.
- We can use **colon notation** to identify groups of matrix entries. $A(i_1\colon i_2, j_1\colon j_2)$ denotes the intersection of the elements between rows i_1 and i_2 and columns j_1 and j_2.
- A **block matrix** is a matrix some of whose entries are themselves matrices.

Example 1.3.3

- Suppose A and B are symmetric (and the same size). Then

$$(A+B)^T = A^T + B^T = A + B,$$

so the sum of symmetric matrices is symmetric.

- Let

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

The diagonal of A consists of the entries 2, 1, 3 and 1. We write

$$\operatorname{diag}(A) = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The columns of A are

$$\mathbf{a}_1 = \begin{bmatrix} 2 \\ 7 \\ 1 \\ -1 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 2 \\ -2 \\ 3 \\ 6 \end{bmatrix}, \quad \mathbf{a}_4 = \begin{bmatrix} -1 \\ 3 \\ 5 \\ 1 \end{bmatrix}.$$

Note that, for example, $\mathbf{a}_2 = A(1{:}4, 2)$. We can write $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 \end{bmatrix}$.

- Let $Q = \begin{bmatrix} 2 & 1 \\ 7 & 1 \end{bmatrix}$, $R = \begin{bmatrix} 2 & -1 \\ -2 & 3 \end{bmatrix}$, $S = \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}$, $T = \begin{bmatrix} 3 & 5 \\ 6 & 1 \end{bmatrix}$.

Then $A = \begin{bmatrix} Q & R \\ S & T \end{bmatrix}$ and $Q = A(1{:}2, 1{:}2)$, $S = A(3{:}4, 1{:}2)$.

1.3.1 Multiplying Matrices

The product of two matrices can be formed in a number of ways but the standard definition of matrix multiplication follows on naturally from the scalar product. It may not seem to be the most natural definition of multiplication, but it is intimately tied to the links we'll develop between matrix algebra and Euclidean geometry.

Definition 13.2 *If A is an $m \times n$ matrix and B is an $n \times p$ matrix then their product C, written $C = AB$ or $C = A \times B$ is the $m \times p$ matrix whose (i,j)-th element is given by the formula*

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}. \tag{1.3.1}$$

Remarks

1. Notice that we can only form AB if B has as many rows as A has columns.
2. If $m \neq p$ we cannot form BA. This indicates the fact that, unlike multiplication of scalars, in general **matrix multiplication is not commutative**.
3. Compare (1.3.1) with (1.2.3). The formula for the (i, j)-th component of C is a scalar product.
4. An alternative definition of matrix multiplication is to use the scalar product directly: the (i, j)-th element of $C = AB$ is the scalar product of the (transposed) i-th row of A with the j-th column of B, that is $c_{ij} = \mathbf{a}_i.\mathbf{b}_j$.
5. If A and B are of the same size, an obvious product would be formed by letting C be the matrix whose (i, j)-th entry is $a_{ij}b_{ij}$. This componentwise product is known as the **Hadamard product**. It has many fewer applications than matrix multiplication.
6. If A is a square matrix then we can form matrix powers in an analogous fashion to the powers of scalars: $A^2 = A \times A$, $A^3 = A \times A \times A, \ldots$. All the powers will be square.
7. Recall that the k-th row of a matrix forms the k-th column of its transpose. Since the (j, i)-th entry of AB is formed from the scalar product of the j-th row of A with the i-th column of B it follows that $(AB)^T = B^T A^T$.

Examples 1.3.4

(i) $\begin{bmatrix} 1 & 4 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 \times 2 + 4 \times 3 & 1 \times (-1) + 4 \times 0 \\ 0 \times 2 + 6 \times 3 & 0 \times (-1) + 6 \times 0 \end{bmatrix} = \begin{bmatrix} 14 & -1 \\ 18 & 0 \end{bmatrix}.$

Let's view this in terms of row and column interactions.

$$\begin{bmatrix} 1\ 4 \\ 0\ 6 \end{bmatrix} \begin{bmatrix} 2 & \text{-1} \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 14 & -1 \\ 18 & 0 \end{bmatrix}$$

(ii) $\begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} 2 \times 1 + (-1) \times 0 & 2 \times 4 + (-1) \times 6 \\ 3 \times 1 + 0 \times 0 & 3 \times 4 + 0 \times 6 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 3 & 12 \end{bmatrix}.$

(iii) $AB = \begin{bmatrix} 2 & 6 & -3 \\ 1 & -9 & 2 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 3 \\ 6 & 2 \end{bmatrix} = \begin{bmatrix} -16 & 14 \\ 13 & -22 \\ 9 & 5 \end{bmatrix}.$

One cannot form BA, but $B^T A = \begin{bmatrix} 1 & 0 & 6 \\ 1 & 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & 6 & -3 \\ 1 & -9 & 2 \\ 3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 20 & 6 & 3 \\ 11 & -21 & 5 \end{bmatrix}.$

Furthermore, $B^T A^T = \begin{bmatrix} 1 & 0 & 6 \\ 1 & 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 & 3 \\ 6 & -9 & 0 \\ -3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} -16 & 13 & 9 \\ 14 & -22 & 5 \end{bmatrix} = (AB)^T.$

(iv) Let

$$C = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}.$$

Then

$$C^2 = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 9 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}, \quad C^3 = \begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 27 & 0 \\ 0 & 0 & 0 & -8, \end{bmatrix}.$$

Let

$$D = \begin{bmatrix} -5 & 0 & 0 & 0 \\ 0 & 11 & 0 & 0 \\ 0 & 0 & -8 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}.$$

Then

$$CD = \begin{bmatrix} -10 & 0 & 0 & 0 \\ 0 & 11 & 0 & 0 \\ 0 & 0 & -24 & 0 \\ 0 & 0 & 0 & -12 \end{bmatrix} = DC.$$

Multiplication of diagonal matrices is componentwise and is commutative.

Examples 1.3.5

Definition 13.3 *The* **identity matrix** *of dimension n is the diagonal matrix whose diagonal entries are all equal to one. We write* $I = \mathrm{diag}(1, 1, \ldots, 1)$*. If we want to make the dimension clear we write* I_n *to denote the* $n \times n$ *identity matrix*

Theorem 13.4 *For any* $m \times n$ *matrix* A, $I_m A = AI_n = A$.

Proof. Exercise. ∎

Note that λI is simply the matrix with λ along its diagonal and zero everywhere else. Let $C = \lambda I_n$ and $B = CA$ then, since $c_{ik} = 0$ unless $i = k$,

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

so $(\lambda I)A = \lambda A$.

Example 1.3.6

Matrix-Vector Product

Since a vector can be viewed as a matrix with one column, matrix multiplication works if we multiply a matrix (with n columns) by a vector (with n elements). The result will be a vector which has as many elements as the matrix has rows and its entries will be the scalar products of the rows of the matrix with the vector. That is, if A is an $m \times n$ matrix and $\mathbf{x}$ is a vector of length n then $\mathbf{b} = A\mathbf{x}$ is a vector of length m and

$$b_i = \sum_{k=1}^{n} a_{ik} x_k = \mathbf{a}_i \cdot \mathbf{x}.$$

Examples 1.3.7

(i) $$\begin{bmatrix} 4 & 3 & 2 \\ 2 & 1 & 5 \\ 0 & 6 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 9 \\ 1 \end{bmatrix}.$$

(ii) $$\begin{bmatrix} -1 & 7 & -2 \\ 3 & 1 & 0 \\ 2 & 1 & -5 \\ 1 & 0 & 10 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

Multiplying a matrix by a column of I picks out a column of the matrix.

(iii) Writing a matrix in terms of its columns, (1.3.1) gives

$$Ax = \left[\ \mathbf{a}_1 \ \middle| \ \mathbf{a}_2 \ \middle| \ \cdots \ \middle| \ \mathbf{a}_n \ \right] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n.$$

A matrix-vector product is a linear combination of the columns of the matrix.

(iv) We can premultiply matrices by row vectors. The result is a row vector.

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 5 \end{bmatrix}.$$

Example 1.3.8

- We cannot multiply column vectors together but we can multiply row and column vectors of the same length. Let

$$\mathbf{x} = \begin{bmatrix} -2 \\ 1 \\ -3 \end{bmatrix}, \ \mathbf{y} = \begin{bmatrix} 4 \\ 5 \\ 2 \end{bmatrix}.$$

Then $\mathbf{x}^T\mathbf{y} = (-2) \times 4 + 1 \times 5 + (-3) \times 2 = -9 = \mathbf{x}.\mathbf{y}$.

This result is true in general. That is, $\mathbf{x}^T\mathbf{y} = \mathbf{x}.\mathbf{y}$.

- We can multiply a column vector by a row vector, too. This is called the **outer product** of two vectors (recall that the scalar product is also known as the inner product). The result is a matrix. Using the vectors from the last example,

$$\mathbf{x}\mathbf{y}^T = \begin{bmatrix} -2 \\ 1 \\ -3 \end{bmatrix} \begin{bmatrix} 4 & 5 & 2 \end{bmatrix} = \begin{bmatrix} -8 & -10 & -4 \\ 4 & 5 & 2 \\ -12 & -15 & -6 \end{bmatrix}.$$

Notice that all the rows of $\mathbf{x}\mathbf{y}^T$ are multiples of $\mathbf{y}^T$ and every column is a multiple of $\mathbf{x}$.

1.3.2 Geometry of Matrix-Vector Products

Hopefully, the reason for choosing the definition of matrix multiplication used in (1.3.1) will become clear as the course progresses. It has to do with the role of matrices in carrying out **linear transformations**. We will see this by the action of matrices on vectors in Euclidean geometry.

Example 1.3.9

- Let $\mathbf{u} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ and $S = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. Evaluate $R\mathbf{v}$, $\mathbf{w} = R\mathbf{u}$, $S\mathbf{v}$, $\mathbf{x} = S\mathbf{u}$.

  Sketch $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ and $\mathbf{x}$ and describe what you see.

$$\begin{aligned} R\mathbf{v} &= \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \mathbf{u}, \quad \mathbf{w} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}, \\ S\mathbf{v} &= \begin{bmatrix} -2 \\ -1 \end{bmatrix} = \mathbf{w}, \quad \mathbf{x} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \end{aligned}$$

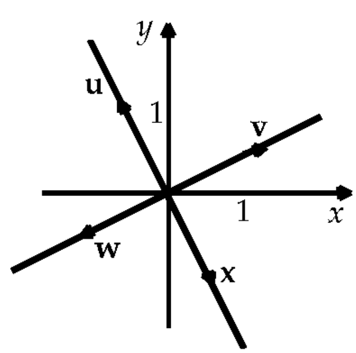


Figure 1.11: Action of matrices.

From Figure 1.11 we see that the action of R is to rotate a vector anticlockwise through a right angle. The action of S is to rotate a vector through a half circle. Furthermore,

$$R^2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = S, \ \ S^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

Hence R^2 and S^2 can be used to perform rotations of double the size of the individual rotations (interpreting I as a rotation through a full circle). Further calculations show that multiplying together copies of R and S will give us matrices whose action composes several rotations together. It is only through the definition of matrix multiplication we have chosen that this is true. We will return to this idea in detail later.

1.3.3 Distributivity and Associativity

We remarked earlier that matrix multiplication is not a commutative operation. That is, AB does not necessarily equal BA. However matrix multiplication and addition do inherit some important properties from scalar operations. Here is a list for matrices which conform (i.e., are of the right size for the operations to be defined)

1. $A + B = B + A$ (commutativity), $A + (B + C) = (A + B) + C$ (associativity).
2. $A(B + C) = AB + AC$ and $(B + C)A = BA + CA$ (distributive laws).
3. $A(BC) = (AB)C$ (associativity).
4. If b and c are scalars $A(cB) = c(AB)$ and $c(bA) = cb(A)$ (homogeneity).

Most of these results follow straight from the definitions of matrix multiplication and addition. For example, let $X = A(B + C)$ and $Y = AB + AC$, then

$$x_{ij} = \sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj}) = \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj} = y_{ij},$$

where A has n columns and B and C have n rows.

The results on homogeneity follow from associativity by letting $B = bI$ and $C = cI$.

So long as the dimensions conform, matrices can be replaced by vectors in the above identities.

Examples 1.3.10

(i) The distributive law holds when B and C are vectors: $A(\mathbf{b} + \mathbf{c}) = A\mathbf{b} + A\mathbf{c}$. For example

$$\begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 20 \\ 15 \end{bmatrix}, \ \ \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \ \ \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 16 \\ 14 \end{bmatrix}.$$

(ii) The product of outer products $(\mathbf{x}\mathbf{u}^T)(\mathbf{v}\mathbf{w}^T)$ is an outer product. Suppose that $\mathbf{u}^T\mathbf{v} = a$. Then

$$(\mathbf{x}\mathbf{u}^T)(\mathbf{v}\mathbf{w}^T) = \mathbf{x}(\mathbf{u}^T\mathbf{v})\mathbf{w}^T = \mathbf{x}a\mathbf{w}^T = a\mathbf{x}\mathbf{w}^T.$$

1.3.4 Inverse Matrices

So far we have seen how to add, subtract and multiply together matrices. A natural question to ask is how do we divide one matrix by another? The notation $A \div B$ is not used. Instead, notice that if $a \div b = a \times (1/b)$. In other words, division by b is the same as multiplication by the reciprocal of b. The reciprocal is an example of an **inverse**. You will have seen inverses in the context of functions, too. The reciprocal is the unique number x which satisfies $bx = 1$ and exists so long as $b \neq 0$. It is the **multiplicative inverse of a number**. When working with matrices, we use the **matrix inverse**.

Definition 13.5 *For a given matrix A, the inverse matrix, denoted A^{-1}, is the matrix for which $A^{-1}A = AA^{-1} = I$.*

Notice that since matrix multiplication is not commutative we have to insist that both $A^{-1}A$ and AA^{-1} equal the identity matrix. The inverse matrix does not always exist. It can only exist if A is square, otherwise $A^{-1}A$ and AA^{-1} would be of different sizes, and it doesn't always exist for square matrices. If the inverse of A does exist then it is unique and A is called an **invertible matrix**. So instead of writing $A \div B$ we write AB^{-1} **or** B^{-1}A. A square matrix which has no inverse is known as a **singular matrix**. Invertible matrices are also known as **nonsingular matrices**.

Examples 1.3.11

(i) $\begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} -5+6 & -10+10 \\ 3-3 & 6-5 \end{bmatrix} = I$ and

$\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} -5+6 & 2-2 \\ -15+15 & 6-5 \end{bmatrix} = I.$

So $\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}^{-1} = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}$ and $\begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$.

(ii) If $B = A^{-1}$ then $BA = AB = I$ and so $B^{-1} = A$ or $\left(A^{-1}\right)^{-1} = A$, which generalises the result above.

(iii) If A is square and $AB = I$ then it must be the case that $BA = I$. We omit a proof, but it means that we don't need to test a candidate inverse for A on both sides.

(iv) $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} da-bc & db-bd \\ -ca+ac & -cb+ad \end{bmatrix} = \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix}$.

The last example can be used as the proof of the following theorem for 2×2 matrices.

Theorem 13.6 *The inverse of $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ exists if and only if $ad \neq bc$ in which case*

$$A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

This is a very useful formula. Formulae exist for bigger matrices but they are very messy and of little practical use in general. However, for certain matrices much simpler formulae apply. For example if A is a diagonal matrix none of whose diagonal entries are zero, it is invertible and its inverse is the diagonal matrix whose (i, i)-th entry is $1/a_{ii}$ (the proof is an exercise).

Examples 1.3.12

Exercises: Matrices

For questions 1–8, let $A = \begin{bmatrix} 1 & 3 \\ 5 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 3 \\ -2 & -1 \end{bmatrix}$, $C = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}$, $D = \begin{bmatrix} -1 & 2 & -2 \\ -3 & 0 & 2 \end{bmatrix}$,

$E = \begin{bmatrix} -3 & 2 & 1 \\ -1 & 1 & 0 \\ -2 & 1 & 1 \end{bmatrix}$ and $F = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 0 & -1 \\ 2 & -1 & 0 \\ -1 & 4 & 5 \end{bmatrix}$.

1. Find A^T, D^T and F^T.

2. Find $A + B$, $3A - 2B$, $A + B^T + 3C$ and $E - E^T$.

3. Write down the matrices $D(1, 1{:}3)$, $E(2{:}3, 1{:}3)$ and $F(2{:}3, 1{:}3)$.

4. Write down the block matrices $G = \begin{bmatrix} A \\ B \end{bmatrix}$, $\begin{bmatrix} D \\ E \end{bmatrix}$, $\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$ and $\begin{bmatrix} G & F \end{bmatrix}$.

5. Find AB, CD, $D^T B$ and FE.

6. Find A^2, C^2, C^3 and C^4. Can you give a formula for C^k?

7. Find $\operatorname{diag}(A)$, $\operatorname{diag}(B)$, $\operatorname{diag}(C^3)$.

8. Find $A^T A$, AA^T, FF^T and $F^T F$.

9. Use the definition of the transpose to show that for any matrix P, $(P^T)^T = P$ and for any matrices P and Q of the same size, $(P + Q)^T = P^T + Q^T$.

10. Indicate whether the following statements are true or false. Justify your answers.

    (a) If A^2 is defined then A is necessarily square.

    (b) If AB and BA are defined then A and B are square.

    (c) If the first two columns of B are identical then so are the first two columns of AB.

    (d) If the first two rows of B are identical then so are the first two rows of AB.

11. Suppose that A is an $n \times n$ skew-symmetric matrix. Using the definition of the transpose, prove that every diagonal element of A must be a zero.

12. Show that the sum of skew-symmetric matrices is skew-symmetric, too.

13. Which of the following expressions are guaranteed to equal $(A - B)^2$ for all square matrices A and B of the same size? Justify your answer and give counterexamples for those which don't.

    (a) $A^2 - B^2$ (b) $(B - A)^2$ (c) $A(A - B) - B(A - B)$ (d) $A^2 - 2AB + B^2$ (e) $A^2 - AB - BA + B^2$

14. Find the 5×5 matrices A, B and C whose entries satisfy the following definitions.

    (a) $a_{ij} = |i - j|$ (b) $b_{ij} = \dfrac{1}{i + j}$ (c) $c_{ij} = \begin{cases} 2^{i-j}, & i \geq j, \\ 0, & i < j. \end{cases}$

15. Using (1.3.1), prove that for any $m \times n$ matrix A, $I_m A = A I_n = A$.

16. Let $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$. Sketch the triangle with vertices $O(0,0)$, $U(1,0)$, $V(0,2)$.

    For the following matrices and vectors, compute $\mathbf{p} = A\mathbf{0} + \mathbf{b}$, $\mathbf{x} = A\mathbf{u} + \mathbf{b}$, $\mathbf{y} = A\mathbf{v} + \mathbf{b}$ and sketch the triangle PXY whose vertices have position vectors $\mathbf{p}$, $\mathbf{x}$ and $\mathbf{y}$.

    (a) $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

    (b) $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

    (c) $A = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

    (d) $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

    (e) $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

    (f) $A = \begin{bmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

17. Show that if $\mathbf{u}$ and $\mathbf{v}$ are two vectors of the same length and $A = \mathbf{u}\mathbf{v}^T$ then $A^2 = (\mathbf{v} \cdot \mathbf{u})A$ and hence (by induction, or otherwise) that $A^k = (\mathbf{v} \cdot \mathbf{u})^{k-1} A$.

    Compute $\left(\mathbf{u}\mathbf{v}^T\right)^k$ when $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

    Compare your answer with that for C^k in question 6.

18. How many 3×3 matrices A can you find such that

$$A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y \\ 2x + y \\ 0 \end{bmatrix}$$

    for all choices of x, y and z?

19. Let A be an $n \times n$ diagonal matrix and let B be the diagonal matrix whose diagonal entries are $b_{ii} = 1/a_{ii}$. Show that $AB = I$. You may assume that all the diagonal entries of A are nonzero.

20. Find the inverses of the following matrices.

    (a) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{bmatrix}$ (e) $\frac{1}{25}\begin{bmatrix} -7 & 24 \\ 24 & 7 \end{bmatrix}$

1.4 Summary

Here is a list of skills you are expected to have picked up from this chapter.

1. Understand all terms in **bold face** (all sections).
2. Understand all the numbered definitions and statement of all numbered theorems and recognise the relevance of the numbered equations (all sections).
3. Be able to describe geometric operations needed to transform congruent shapes into each other (§1.1 Exercises).
4. Be able to draw a curve given a parametric expression for it (§1.1).
5. Be able to convert relationships expressed in Cartesian coordinates to polar form and vice versa (§1.1).
6. Use vector algebra to manipulate vectors in Cartesian form (§1.2).
7. Measure the length of vectors and find unit vectors in any given direction (§1.2).
8. Carry out simple problems, such as resolving angles, using the scalar product (§1.2).
9. Manipulate matrices using matrix algebra (§1.3).
10. Understand when two matrices can be added or multiplied together (§1.3).
11. Recognise the geometric interpretation of the matrix-vector product (§1.3)
12. Be able to find the inverse of a 2×2 matrix, if it exists (§1.3).