
Learning from Data

Supplementary Mathematics (Vector and Linear Algebra)

David Barber

1 An Introduction to Vectors

We are all familiar with the fact that if B is two miles from A, and if C is two miles from B, then C is not necessarily four miles from A. Only in very special circumstances are distances compounded according to the ordinary arithmetical law of addition. Actually, there are many other entities which behave in this way as distances rather than as ordinary numbers; this motivates the study of the algebra of vectors, along with associated basic geometrical concepts.

Scalars

A scalar is a quantity which has magnitude but which is not related to any definite direction in space. A scalar is completely specified by a number.

Vectors

A vector is an entity which obeys the same law of addition as a distance does. It has magnitude and is also related to a definite direction in space. It follows that any vector may be represented by a straight line whose direction is that of the vector and whose length represents the magnitude of the vector according to a convenient scale.

There are several ways of denoting vectors. In these notes, we shall use boldface, *e.g.*, \mathbf{r} . Other textbooks, particularly continental, use \vec{r} . When hand written, one commonly uses the underscore notation \underline{r} .

A vector of zero magnitude is called the zero vector and is denoted by $\mathbf{0}$. The magnitude of a vector is called its length. It is often convenient to denote the length of a vector \mathbf{r} by r . A similar convention is used when vectors are denoted by other letters. An alternative notation for the length of a vector \mathbf{r} , is $|\mathbf{r}|$. The length of a vector is also called its *norm*.

To signify a vector of unit length, we use the notation, $\hat{\mathbf{r}}$, which signifies that $|\mathbf{r}| = 1$.

1.1 Vector Addition

Two vectors are added in the same way as distances, so we define the addition of vectors by the parallelogram law:

If \mathbf{a} is represented by the directed line \overline{OA} and \mathbf{b} is represented by \overline{OB} , the $\mathbf{a} + \mathbf{b}$ is defined as the vector represented by \overline{OC} where OACB is the

completed parallelogram fig(1). \overline{OC} is the sum or resultant of \overline{OA} and \overline{OB} . \overline{AC} has the same magnitude and direction as \overline{OB} so that we may alternatively choose \overline{AC} to represent \mathbf{b} . Thus, the resultant of \overline{OA} and \overline{AC} is \overline{OC} . If $\mathbf{a} + \mathbf{b} = \mathbf{0}$, i.e., if O and C coincide, then $\mathbf{b} = -\mathbf{a}$. Thus $-\mathbf{a}$ is a vector which has the same length as \mathbf{a} but in the opposite direction. If \mathbf{a} is represented by \overline{OA} , then $-\mathbf{a}$ is represented by \overline{AO} .

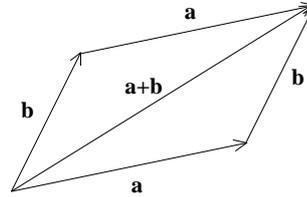


Figure 1:

It is evident from the definition that the commutative law holds

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a} \quad (1)$$

as does the associative law

$$\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + \mathbf{b} + \mathbf{c}, \quad (2)$$

The vector $\mathbf{a} + \mathbf{a}$ is naturally called $2\mathbf{a}$ and is a vector in the same direction as \mathbf{a} but of twice its length. Similarly, $m\mathbf{a}$ is a vector in the direction as \mathbf{a} but of length ma .

Evidently,

$$m(n\mathbf{a}) = n(m\mathbf{a}) = nma, \quad (3)$$

and

$$(m + n)\mathbf{a} = m\mathbf{a} + n\mathbf{a}. \quad (4)$$

Also, from the similar triangles in fig(2)¹, we see that

$$m(\mathbf{a} + \mathbf{b}) = m\mathbf{a} + m\mathbf{b} \quad (5)$$

If a vector \mathbf{r} can be represented as a sum of vectors $\mathbf{a} + \mathbf{b} + \dots + \mathbf{d}$, we say that \mathbf{r} can be resolved into components $\mathbf{a}, \mathbf{b}, \dots, \mathbf{d}$. It is important to realise that any 3 dimensional vector can be resolved into 3 components in *any* 3 given directions which are not coplanar.

Basis Vectors

A set of vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ is a basis for the n -dimensional space provided that the set of vectors is linearly independent. By linearly independent we mean that no vector in the set can be expressed as a linear combination of the others.

Position Vectors

The position of a point P in 3 dimensional space is described by its coordinates, represented by the ordered tuple, (x, y, z) (or, in two dimensions, the pair (x, y)). The point P can also be represented by the position vector \mathbf{p} ,

¹ Note that I will sometimes place the arrow representing the direction of the vector at the end of the vector, and sometimes in the middle. There is no significance in this.

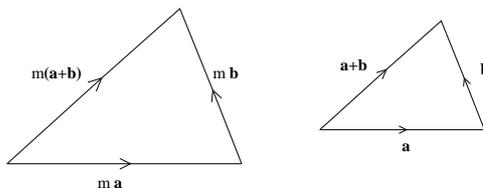


Figure 2:

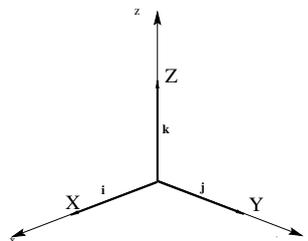


Figure 3:

which is a vector whose initial point is the origin of the coordinate system (usually the point $(0, 0, 0)$) and whose terminal point is P .

We use the same notation for vectors and position vectors, and their distinction will sometimes be blurred – however, it will always be clear from the context which type of vector we have in mind.

1.2 Component Notation for Vectors

It is often convenient to fix the component directions of vectors from the outset, and define all vectors according to these directions.

One convenient set of basis vectors is given by those unit vectors parallel with the co-ordinate axes. See fig(3).

An alternative notation for $\mathbf{r} = \alpha\mathbf{i} + \beta\mathbf{j} + \gamma\mathbf{k}$ is $\mathbf{r} = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$. This is called a column vector (more correctly we should call this the column representation of the vector). It is the transpose of the row vector, (α, β, γ) , so that $\mathbf{r} = (\alpha, \beta, \gamma)^T$, where T denotes the transpose operator.

Some Special Vectors

The unit vectors along the direction of the coordinate axes X, Y, Z are commonly denoted as

$$\mathbf{i} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{j} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{6}$$

respectively. Similarly, in two dimensions, $\mathbf{i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Example 1 Find the summation of the vectors \mathbf{a} and \mathbf{b} , defined as (see fig(4))

$$\mathbf{a} = \begin{pmatrix} 6 \\ 2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} \tag{7}$$

Solution

$$\mathbf{a} + \mathbf{b} = \begin{pmatrix} 6+3 \\ 2+6 \end{pmatrix} = \begin{pmatrix} 9 \\ 8 \end{pmatrix} \quad (8)$$

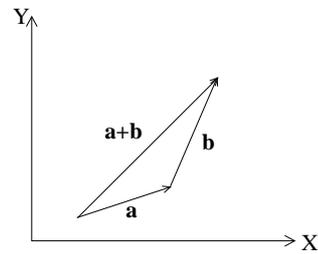


Figure 4:

► **Exercise 1** The position vectors of the four points A, B, C, D are $\mathbf{a}, \mathbf{b}, 2\mathbf{a} + 3\mathbf{b}$, and $\mathbf{a} - 2\mathbf{b}$ respectively. Express AC, DB, BC and CD in terms of \mathbf{a} and \mathbf{b} .

► **Exercise 2** Find the vector that has initial point $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and terminal point $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Likewise, for the initial point $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ and terminal point $\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$.

1.3 Higher Dimensional Vectors

Although a little difficult to visualise, vectors of any (natural) dimension can be defined. Such concepts are extremely common, for example in dealing with data bases. For example, a 4 dimensional vector, could be written as

$$\mathbf{r} = \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} = (\alpha, \beta, \gamma, \delta)^T. \quad (9)$$

The concepts that we shall develop generally hold for higher dimensional vectors.

► **Exercise 3** Let $\mathbf{a} = (2, 1, 0)^T$, $\mathbf{b} = (-2, 1, 0)^T$, $\mathbf{c} = (0, 2, 1)^T$. Find the components of the vectors

$$3\mathbf{a} + 2\mathbf{b} - 7\mathbf{c}, \quad 5\mathbf{a} + 4\mathbf{b}, \quad 3\mathbf{b} - 7\mathbf{c} + \alpha\mathbf{a} \quad (10)$$

► **Exercise 4** Let $\mathbf{a} = (2, 1, 0)^T$, $\mathbf{b} = (1, 1, 2)^T$, $\mathbf{c} = (-3, 2, 1)^T$. Find $\alpha, \beta, \gamma \in \mathcal{R}$ such that

$$(1, 1, 1)^T = \alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} \quad (11)$$

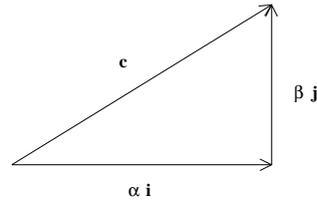


Figure 5:

1.4 The Length of a Vector

In two dimensions, the length of the vector $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ is $\sqrt{x_1^2 + x_2^2}$, which is an application of Pythagoras' theorem.

How can we find the length of a vector in three dimensions? Let's write the vector $\mathbf{x} = (x_1, x_2, x_3)^T = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$. We can apply (directly) Pythagoras' theorem to vectors which consist of two perpendicular components. To achieve this, define a new vector \mathbf{l} with length γ in direction $\hat{\mathbf{l}}$. $\gamma\hat{\mathbf{l}} = x_1\mathbf{i} + x_2\mathbf{j}$. With this definition $\mathbf{x} = \gamma\hat{\mathbf{l}} + x_3\mathbf{k}$. The vector $\hat{\mathbf{l}}$ lies in the plane XY , since it is a summation of the \mathbf{i} and \mathbf{j} vectors, and is therefore orthogonal to \mathbf{k} . We can therefore apply Pythagoras' theorem to this vector to obtain

$$|\mathbf{x}|^2 = \gamma^2 + x_3^2 \quad (12)$$

Now γ is the length of the vector $x_1\mathbf{i} + x_2\mathbf{j}$. Again, we can apply Pythagoras' theorem to obtain $\gamma^2 = x_1^2 + x_2^2$. Putting this all together, we obtain, for the length of the vector $\mathbf{x} = (x_1, x_2, x_3)^T$,

$$|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + x_3^2} \quad (13)$$

Similarly, we define the lengths (or norms) of a vector \mathbf{x} of dimension n as

$$|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \quad (14)$$

Example 2 A ship P is travelling due East at 30 kilometres per hour and a ship Q is travelling due South at 40 kilometres per hour. Both ships keep constant speed and course. At noon they are each 10 km from the point of intersection, O , of their courses, and moving towards O (the ships do not collide!). Find the coordinates, with respect to axes Ox eastwards and Oy northwards, of P and Q at time t hours past noon, and find the distance PQ at this time. Find the time at which P and Q are closest to one another. Find the magnitude and direction of the velocity of Q relative to P , indicating the direction on a diagram. Show that, at the position of closest approach, the bearing of Q from P is θ degrees South of East, where $\tan \theta = \frac{3}{4}$.

Solution The velocity vector of ship P is $(30, 0)^T$, and of ship Q is $(0, -40)^T$, as depicted in fig(6). At noon, the two ships are at the points P, Q as depicted in the figure. From that time onwards, their position vectors \mathbf{p} and \mathbf{q} are

$$\mathbf{p} = \begin{pmatrix} -10 + 30t \\ 0 \end{pmatrix}, \quad \mathbf{q} = \begin{pmatrix} 0 \\ 10 - 40t \end{pmatrix} \quad (15)$$

Their relative motion \mathbf{r} is given by

$$\mathbf{r} = \mathbf{p} - \mathbf{q} = \begin{pmatrix} -10 + 30t \\ -10 + 40t \end{pmatrix} \quad (16)$$

The length of this vector is

$$|\mathbf{r}|^2 = 100 - 600t + 900t^2 + 100 - 800t + 1600t^2 \quad (17)$$

$$= 200 - 1400t + 2500t^2 \quad (18)$$

$$= 100(25t^2 - 14t + 2) \quad (19)$$

$$= 100 \left(\left(5t - \frac{14}{10} \right)^2 - \left(\frac{14}{10} \right)^2 + 2 \right) \text{ Completing the square} \quad (20)$$

Thus the minimal distance between the ships occurs as $t = \frac{14}{50}$ hours past noon. At this time, the relative motion of the ships is given by

$$\mathbf{r} = \begin{pmatrix} -10 + \frac{3 \cdot 14}{5} \\ -10 + \frac{4 \cdot 14}{5} \end{pmatrix} = \frac{1}{5} \begin{pmatrix} -50 + 42 \\ -50 + 56 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} -8 \\ 6 \end{pmatrix}. \quad (21)$$

The angle this vector makes with the eastwards axis is $\tan \theta = 6/8 = 3/4$.

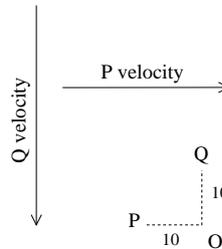


Figure 6:

- **Exercise 5** For which $\alpha \in \mathcal{R}$ does

$$\alpha(1, 1, 1)^T - (3, 2, 4)^T \quad (22)$$

have length $\sqrt{6}$?

- **Exercise 6** A boat sails for one hour at 4 km/hr (relative to the water) on a steady compass heading of 40° east of north. The boat is simultaneously carried along by a current. At the end of the hour, the boat is 6.12 km from its starting point. The line from its starting point to its location lies 60° east of north. Find the easterly and northerly components of the waters velocity.
- **Exercise 7** A triangle is defined by the vertices of the three vectors, \mathbf{a} , \mathbf{b} and \mathbf{c} that extend from the origin. In terms of \mathbf{a} , \mathbf{b} and \mathbf{c} show that the vector sum of the successive sides of the triangle ($AB + BC + CA$) is zero.

1.5 The Scalar Product

We define the scalar product (sometimes also called the inner product) of two vectors \mathbf{a} and \mathbf{b} as $|\mathbf{a}||\mathbf{b}| \cos \theta$ where θ is the angle between the directions of the two vectors.

We write

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta = \mathbf{b} \cdot \mathbf{a} \quad (23)$$

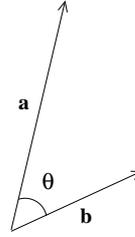


Figure 7: The scalar product of two vectors is a function of the angle between the two vectors and their lengths.

The scalar product of two vectors is not a vector but a scalar, and scalar multiplication is commutative². If two vectors are perpendicular, their scalar product is zero since $\cos \theta = 0$. If two vectors \mathbf{a} and \mathbf{b} are perpendicular, we write $\mathbf{a} \perp \mathbf{b}$. In particular,

$$\mathbf{i} \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{i} = 0. \quad (24)$$

If \mathbf{a} and \mathbf{b} have the same direction, $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}|$. In particular

$$\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} = 1. \quad (25)$$

Since $\mathbf{a} \cdot \mathbf{b}$ is the length of \mathbf{a} multiplied by the projected length of \mathbf{b} on \mathbf{a} , it follows that

$$\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c} \quad (26)$$

for the projection of $\mathbf{b} + \mathbf{c}$ on \mathbf{a} is the sum of the projections of \mathbf{b} and \mathbf{c} on \mathbf{a} . Evidently,

$$m(\mathbf{a} \cdot \mathbf{b}) = (m\mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (m\mathbf{b}). \quad (27)$$

How to evaluate the Scalar Product

From above properties of the scalar product, we see that if

$$\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}, \quad \mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k} \quad (28)$$

then

$$\mathbf{a} \cdot \mathbf{b} = (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \cdot (b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}) \quad (29)$$

$$= a_1b_1\mathbf{i} \cdot \mathbf{i} + a_1b_2\mathbf{i} \cdot \mathbf{j} + a_1b_3\mathbf{i} \cdot \mathbf{k} + a_2b_1\mathbf{j} \cdot \mathbf{i} + a_2b_2\mathbf{j} \cdot \mathbf{j} + a_2b_3\mathbf{j} \cdot \mathbf{k} + a_3b_1\mathbf{k} \cdot \mathbf{i} + a_3b_2\mathbf{k} \cdot \mathbf{j} + a_3b_3\mathbf{k} \cdot \mathbf{k} \quad (30)$$

Due to the orthogonality of the chosen basis vectors, this gives

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3 \quad (31)$$

This generalises to any dimensional vectors, which enables us to also think of angles between vectors in higher dimensional spaces. In general, the angle between two vectors \mathbf{a} and \mathbf{b} is then found from

$$\cos \theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|} \quad (32)$$

It follows from Pythagoras' theorem that the length of a vector is related to the scalar product as

² An operator \otimes is *commutative* if $A \otimes B \equiv B \otimes A$.

$$\mathbf{a} \cdot \mathbf{a} = a_1^2 + a_2^2 + a_3^2 = |\mathbf{a}|^2 \quad (33)$$

Again, this concept generalises to any dimension of vector. For convenience, the notation \mathbf{a}^2 is often used in place of $\mathbf{a} \cdot \mathbf{a}$. Note that the angle between two vectors (regardless of the dimension of the space in which they lie) can be found from evaluating scalar products alone.

Example 3 Find the angle between the vectors $\mathbf{r}_1 = (1, 1, 1)^T$ and $\mathbf{r}_2 = (-1, 2, 3)^T$.

Solution The lengths of the two vectors are $|\mathbf{r}_1| = \sqrt{1+1+1} = \sqrt{3}$ and $|\mathbf{r}_2| = \sqrt{1+4+9} = \sqrt{14}$. Thus

$$\cos \theta = \frac{\mathbf{r}_1 \cdot \mathbf{r}_2}{|\mathbf{r}_1||\mathbf{r}_2|} = \frac{-1+2+3}{\sqrt{42}} = \frac{4}{\sqrt{42}} \quad (34)$$

Example 4 Find the angle between a diagonal of a unit cube and one of its edges.

Solution The 8 vertices of the cube lie at the 8 points $b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, where $b_i = 0$ or 1 . The diagonal from the origin to the opposite vertex is the position $(1, 1, 1)^T$. One of the edges of the cube is $(1, 0, 0)^T$. The angle between the diagonal and this X -edge is therefore

$$\cos \theta = \frac{1+0+0}{\sqrt{1+1+1}} = \frac{1}{\sqrt{3}} \quad (35)$$

This gives $\theta = \cos^{-1}(1/\sqrt{3}) \approx 54^\circ 44'$

- **Exercise 8** Using vector algebra, show that, for the triangle depicted in fig(8), $c^2 = a^2 + b^2 - 2ab \cos \theta$ where a, b, c are the lengths of the sides of the triangle. (This is a generalisation of Pythagoras' rule to non-right-angled triangles).

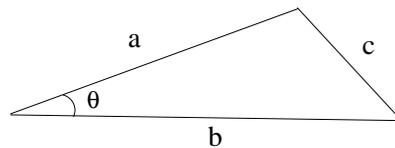


Figure 8:

- **Exercise 9** Find the angles and side lengths of the triangle whose vertices are at $(1, 1, -1)$, $(2, -1, 1)$ and $(-1, 1, 1)$.
- **Exercise 10** A pipe comes diagonally down the south wall of a building, making an angle of 45° with the horizontal. Coming into a corner, the pipe turns and continues diagonally down a west-facing wall, still making an angle of 45° with the horizontal. What is the angle between the south-wall and west-wall sections of the pipe?

Turning a vector into a unit vector If \mathbf{f} is any vector, then $\frac{\mathbf{f}}{|\mathbf{f}|}$ is a unit vector in the direction of \mathbf{f} .

Why? It is clear that dividing by a scalar, $|\mathbf{f}|$ does not change the direction of \mathbf{f} . The squared length of the new vector is

$$\frac{\mathbf{f}}{|\mathbf{f}|} \cdot \frac{\mathbf{f}}{|\mathbf{f}|} = \frac{\mathbf{f} \cdot \mathbf{f}}{|\mathbf{f}|^2} = 1 \quad (36)$$

► **Exercise 11** Establish the identities

$$|\mathbf{u} + \mathbf{v}|^2 + |\mathbf{u} - \mathbf{v}|^2 = 2|\mathbf{u}|^2 + 2|\mathbf{v}|^2 \quad (37)$$

$$\mathbf{u} \cdot \mathbf{v} = \frac{1}{4}|\mathbf{u} + \mathbf{v}|^2 - \frac{1}{4}|\mathbf{u} - \mathbf{v}|^2 \quad (38)$$

► **Exercise 12** Find the angle between a diagonal of a cube and one of its faces.

1.6 The Scalar Product as a Projection Operator

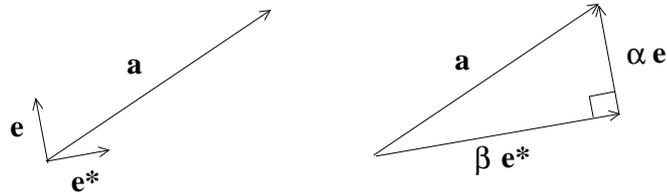


Figure 9: Resolving a vector \mathbf{a} into components along the orthogonal directions \mathbf{e} and \mathbf{e}^* . The projection of \mathbf{a} onto these two directions are lengths α and β along the directions \mathbf{e} and \mathbf{e}^* .

Suppose that we wish to resolve the vector \mathbf{a} into its components along the orthogonal directions specified by the unit vectors \mathbf{e} and \mathbf{e}^* . That is $|\mathbf{e}| = |\mathbf{e}^*| = 1$ and $\mathbf{e} \cdot \mathbf{e}^* = 0$. This is depicted in fig(9). We are required to find the scalar values α and β such that

$$\mathbf{a} = \alpha \mathbf{e} + \beta \mathbf{e}^* \quad (39)$$

Taking the scalar product of both sides of this equation with respect to \mathbf{e} , and then with respect to \mathbf{e}^* we get

$$\mathbf{a} \cdot \mathbf{e} = \alpha \mathbf{e} \cdot \mathbf{e} + \beta \mathbf{e}^* \cdot \mathbf{e}, \quad \mathbf{a} \cdot \mathbf{e}^* = \alpha \mathbf{e} \cdot \mathbf{e}^* + \beta \mathbf{e}^* \cdot \mathbf{e}^* \quad (40)$$

From the orthogonality and unit lengths of the vectors \mathbf{e} and \mathbf{e}^* , this becomes simply

$$\mathbf{a} \cdot \mathbf{e} = \alpha, \quad \mathbf{a} \cdot \mathbf{e}^* = \beta \quad (41)$$

This means that we can write the vector \mathbf{a} in terms of the orthonormal³ components \mathbf{e} and \mathbf{e}^* as

$$\mathbf{a} = (\mathbf{a} \cdot \mathbf{e}) \mathbf{e} + (\mathbf{a} \cdot \mathbf{e}^*) \mathbf{e}^* \quad (42)$$

One can see therefore that the scalar product between \mathbf{a} and \mathbf{e} “projects” the vector \mathbf{a} onto the (unit) direction \mathbf{e} .

³ A set of vectors is orthonormal if they are mutually orthogonal and have unit length.

We can write the projection of a vector \mathbf{a} onto a direction specified by \mathbf{f} as (by finding the unit vector in the direction of \mathbf{f})

$$\frac{\mathbf{a} \cdot \mathbf{f}}{|\mathbf{f}|^2} \mathbf{f} \quad (43)$$

The projection onto the component orthogonal to \mathbf{f} is then given by

$$\mathbf{f} - \frac{\mathbf{a} \cdot \mathbf{f}}{|\mathbf{f}|^2} \mathbf{f} \quad (44)$$

Example 5 Let $\mathbf{u} = (2, -1, 3)^T$ and $\mathbf{a} = (4, -1, 2)^T$. Find the vector component of \mathbf{u} along \mathbf{a} and the vector component of \mathbf{u} orthogonal to \mathbf{a} .

Solution

$$\mathbf{u} \cdot \mathbf{a} = 2 \cdot 4 + (-1) \cdot (-1) + 3 \cdot 2 = 15 \quad (45)$$

$$|\mathbf{a}|^2 = 4^2 + (-1)^2 + 2^2 = 21 \quad (46)$$

Thus, the vector component of \mathbf{u} along \mathbf{a} is

$$\frac{\mathbf{u} \cdot \mathbf{a}}{|\mathbf{a}|^2} = \frac{15}{21} (4, -1, 2)^T = \left(\frac{20}{7}, -\frac{5}{7}, \frac{10}{7} \right)^T \quad (47)$$

and the vector component of \mathbf{u} orthogonal to \mathbf{a} is

$$\mathbf{u} - \frac{\mathbf{u} \cdot \mathbf{a}}{|\mathbf{a}|^2} = (2, 1, -3)^T - \left(\frac{20}{7}, -\frac{5}{7}, \frac{10}{7} \right)^T = \left(-\frac{6}{7}, -\frac{2}{7}, \frac{11}{7} \right)^T \quad (48)$$

An example application: Projecting (Hyper)cubes

We can use the above results to project onto a space defined by two orthonormal basis vectors objects defined in higher dimensions. For example, imagine that we take the unit 3 dimensional cube, defined by the vertices (“corners”) at the points described by the vectors $(0, 0, 0)^T, (0, 0, 1)^T \dots (1, 1, 1)^T$. For a general N -dimensional hypercube, there are 2^N vertices. An edge exists between any two vertices whose vectors differ by only in one component.

Let’s define two basis vectors. We take $\mathbf{e} = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$ and $\mathbf{e}^* = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$.

We then calculate the projections of each vertex in the cube with these two directions, and plot the projected points in fig(10a). Vertices which were connected by an edge in the original cube are connected together in the projection.

We can do the same for higher dimensional cubes. fig(10b) shows the projection of a 5 dimensional hypercube onto two randomly chosen directions.

► **Exercise 13** Plot the projection of a unit 4 dimensional hypercube onto the directions

$$\mathbf{e} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \text{ and } \mathbf{e}^* = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{pmatrix}.$$

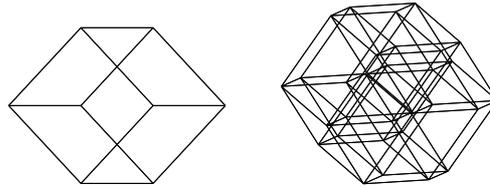


Figure 10: (left) A 3 dimensional unit cube projected onto 2 dimensions. (right) A 5 dim. cube projected onto two (randomly chosen) directions.

- **Exercise 14** How many edges are there in a Hypercube of dimension N ?

Using the scalar product to project higher dimensional objects onto lower dimensional spaces is very common. The visualisation of high dimensional databases is one such common application. However, some care needs to be made in doing this – our intuitions from the 2 or 3 dimensional world can sometimes lead us astray when thinking about higher dimensional objects. You may come across such phenomena in more advanced courses.

The concept of projections of three dimensional objects onto two dimensional surfaces is central to computer graphics.

- **Exercise 15** Find the projection of \mathbf{u} on \mathbf{a}

- (a) $\mathbf{u} = (6, 2)^T$, $\mathbf{a} = (3, -9)^T$
 (b) $\mathbf{u} = (-1, -2)^T$, $\mathbf{a} = (-2, 3)^T$
 (c) $\mathbf{u} = (3, 1, 7)^T$, $\mathbf{a} = (1, 0, 5)^T$
 (d) $\mathbf{u} = (1, 0, 0)^T$, $\mathbf{a} = (4, 3, 8)^T$

- **Exercise 16** For each part of the previous exercise, find the vector component of \mathbf{u} orthogonal to \mathbf{a} .

1.7 Lines and Hyperplanes

A line in 2 (or more) dimensions can be specified as follows. The vector of any point along the line is given, for some s , by the equation

$$\mathbf{p} = \mathbf{a} + s\mathbf{u}, \quad s \in \mathcal{R}. \quad (49)$$

where \mathbf{u} is parallel to the line, and the line passes through the point \mathbf{a} , see fig(11). This is called the parametric representation of the line. An alternative specification can be given by realising that all vectors along the line are orthogonal to the normal of the line, \mathbf{n} . (\mathbf{u} and \mathbf{n} are orthonormal). That is

$$(\mathbf{p} - \mathbf{a}) \cdot \mathbf{n} = 0 \quad (50)$$

or

$$\mathbf{p} \cdot \mathbf{n} = \mathbf{a} \cdot \mathbf{n} \quad (51)$$

If the vector \mathbf{n} is of unit length, the right hand side of the above represents the shortest distance from the origin to the line, drawn by the dashed line in fig(11) (since this is the projection of \mathbf{a} onto the normal direction).

Gradient Vector

The gradient vector is a line which is parallel to the straight line (or, al-

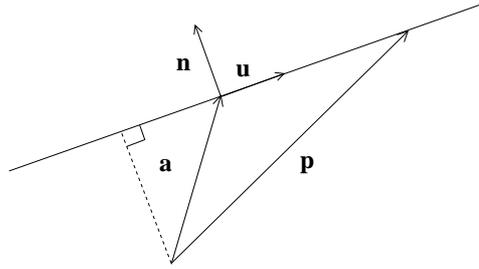


Figure 11: A line can be specified by some position vector on the line, \mathbf{a} , and a unit vector along the direction of the line, \mathbf{u} . In 2 dimensions, there is a unique direction, \mathbf{n} , perpendicular to the line. In three dimensions, the vectors perpendicular to the direction of the line lie in a plane, whose normal vector is in the direction of the line, \mathbf{u} .

ternatively, orthogonal to the normal of the line). If the gradient vector is $(a, b)^T$, the *gradient of a line* is the ratio of the components, $\frac{b}{a}$.

1.8 2 lines in 3 space

As in two dimensions, a line in three dimensions can be specified parametrically in the form

$$\mathbf{p} = \mathbf{a} + t\mathbf{b}, \quad t \in \mathcal{R} \quad (52)$$

where \mathbf{a} is some point on the line, and \mathbf{b} is a vector parallel to the line direction. Note that, whereas in two dimensions, we could also specify a line by a direction orthogonal to the line, in three dimensions, this would specify a plane.

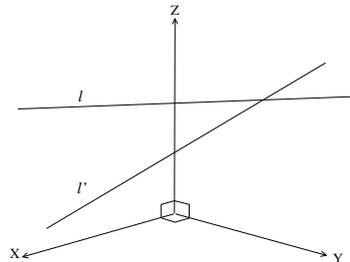


Figure 12: Two lines in three dimensions. In two dimensions, two lines will cross, only if they are parallel. In three dimensions, two lines will cross only if they are coplanar and non-parallel.

► **Exercise 17** $\mathbf{a} = (2, 4, 6)^T$, and $\mathbf{b} = (3, -3, -5)^T$. Compute $\mathbf{a} \cdot \mathbf{b}$. Show that, for any vectors \mathbf{c} and \mathbf{d} , $(\mathbf{c} - \mathbf{d}) \cdot (\mathbf{c} + \mathbf{d}) = |\mathbf{c}|^2 - |\mathbf{d}|^2$. Using this result, compute $(\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} + \mathbf{b})$.

► **Exercise 18** Calculate the minimum distance between the two lines

$$\mathbf{p} = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} + s \begin{pmatrix} 4 \\ 2 \\ 5 \end{pmatrix}, \quad s \in \mathcal{R} \quad (53)$$

and

$$\mathbf{q} = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} + t \begin{pmatrix} -2 \\ 2 \\ 3 \end{pmatrix}, \quad t \in \mathcal{R} \quad (54)$$

1.9 Planes and Hyperplanes

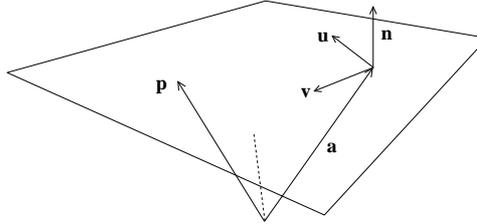


Figure 13: A plane can be specified by a point in the plane, \mathbf{a} and two, non-parallel directions in the plane, \mathbf{u} and \mathbf{v} . The normal to the plane is unique, and in the same direction as the directed line from the origin to the nearest point on the plane.

Mathematically speaking, a line is a one dimensional hyperplane. To specify a 2 dimensional hyperplane (which we will refer to as a “plane”), one way to do this is as follows. Specify two vectors \mathbf{u} and \mathbf{v} that lie in the plane (they need not be mutually orthogonal), and a position vector \mathbf{a} in the plane, see fig(13). Any vector \mathbf{p} in the plane can then be written as

$$\mathbf{p} = \mathbf{a} + s\mathbf{u} + t\mathbf{v}, \quad (s, t) \in \mathcal{R}. \quad (55)$$

An alternative definition is given by considering that any vector within the plane must be orthogonal to the normal of the plane \mathbf{n} .

$$(\mathbf{p} - \mathbf{a}) \cdot \mathbf{n} = 0 \quad (56)$$

or

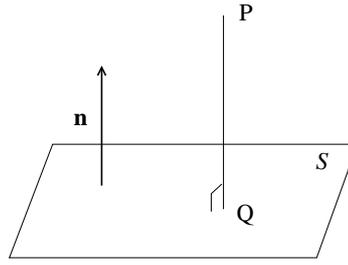
$$\mathbf{p} \cdot \mathbf{n} = \mathbf{a} \cdot \mathbf{n} \quad (57)$$

The right hand side of the above represents the shortest distance from the origin to the plane, drawn by the dashed line in fig(13). The advantage of this representation is that it has the same form as a line. Indeed, this representation of (hyper)planes is independent of the dimension of the space. In addition, only two vectors need to be defined – a point in the plane, \mathbf{a} , and the normal to the plane \mathbf{n} .

1.10 Projection of a point on a plane

Consider a plane S with normal vector $\hat{\mathbf{n}}$ and let P be a point not in S . A point Q in the plane S is called the projection of P on S if $\mathbf{q} - \mathbf{p}$ is parallel to $\hat{\mathbf{n}}$. The distance from P to the plane S is the length $|PQ|$. See fig(14)

Example 6 Given the unit normal to a plane, $\hat{\mathbf{n}}$, and a point in the plane, \mathbf{a} , find an expression for the projection of a point \mathbf{p} onto the plane.

Figure 14: Projecting a point P onto a plane S .

Solution The point \mathbf{p} can be expressed as follows

$$\mathbf{p} = \mathbf{a} + \mathbf{b} + r\hat{\mathbf{n}} \quad (58)$$

where \mathbf{b} is some vector in the plane, orthogonal to $\hat{\mathbf{n}}$, and r is the shortest distance from the point \mathbf{p} to the plane (this must be along the normal). Taking the scalar product of this equation with $\hat{\mathbf{n}}$, we get

$$\mathbf{p} \cdot \hat{\mathbf{n}} = \mathbf{a} \cdot \hat{\mathbf{n}} + r \quad (59)$$

which means that $r = (\mathbf{p} - \mathbf{a}) \cdot \hat{\mathbf{n}}$. The point projected onto the plane is then $\mathbf{a} + \mathbf{b}$ which equals $\mathbf{p} - r\hat{\mathbf{n}}$ from equation (58). Using the value we have found for r , this gives the projection of \mathbf{p} onto the plane as

$$\mathbf{p} - ((\mathbf{p} - \mathbf{a}) \cdot \hat{\mathbf{n}}) \hat{\mathbf{n}}. \quad (60)$$

1.11 Linear Equations

As we shall see, systems of linear equations can be represented by a series of lines, planes (and in general, hyperplanes).

The general equation of a straight line (in 2 dimensions) is

$$ax + by + c = 0 \quad (61)$$

where a, b, c are constants, and x is the distance along the X -axis, and y the distance vertically, along the Y -axis. Another representation of the line is given by

$$y = -\frac{a}{b}x - c \quad (62)$$

where the ratio $-a/b$ represents the “gradient” of the line. In vector notation, equation (61) becomes simply

$$\mathbf{v} \cdot \mathbf{n} + c = 0 \quad (63)$$

where we define the vectors $\mathbf{v} = (x, y)^T$ and $\mathbf{n} = (a, b)^T$. From our previous work, and in particular, equation (51), we know that \mathbf{n} points in the direction perpendicular to the line, since equation (63) is of the form equation (51). Indeed, we know that $-c/|\mathbf{n}|$ is then the shortest distance from the origin to the line (the normal direction \mathbf{n} is not necessarily a unit vector).

Simultaneous linear equations (two dimensions) Consider now two linear equations:

$$a_1x + b_1y + c_1 = 0 \quad (64)$$

$$a_2x + b_2y + c_2 = 0 \quad (65)$$

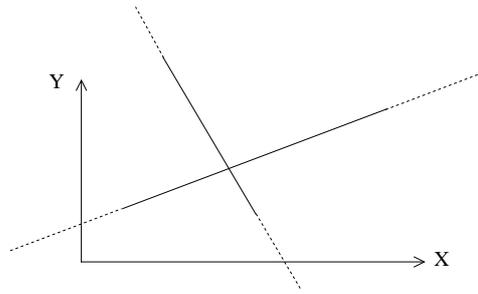


Figure 15:

Each of these equations represents a line in 2-dimensional space. The point at which they cross is the solution of these linear equations. Clearly, they will cross at some point, provided that they are not parallel.

- **Exercise 19** Determine the line of intersection of the planes $x - y + z = 0$ and $2x - 3y - 2z = 1$.
- **Exercise 20** Find the point at which the two lines represented in equation (65) meet. This is called the solution of two simultaneous linear equations. (We shall return to this later, when discussing matrix methods.)
- **Exercise 21** Two planes are defined by the equations

$$\mathbf{x} \cdot \mathbf{a} + \alpha = 0, \quad \mathbf{x} \cdot \mathbf{b} + \beta = 0 \quad (66)$$

Calculate the point x_3 on the second plane where $x_1 = 0, x_2 = 0$. Calculate the nearest point on the first plane to this point, and find this distance.

- Example 7** Find the line of intersection of the planes $x + 2y + 3z = 14$ and $2x + 5y - z = 9$ and find where this line intersects the plane $3x + 8y - 4z = 7$.

Solution First solve the equations

$$x + 2y + 3z = 14 \quad (67)$$

$$2x + 5y - z = 9 \quad (68)$$

The first plus three times the second equation gives $7x + 17y = 41$ or $7x - 41 = -17y$.

The second equation minus two times the first gives $y - 7z = -19$ or $y = 7z - 19$. Lets parameterize the line using $t = y/7$. The line is then gives as

$$x = -17t + \frac{41}{7} \quad (69)$$

$$y = 7t \quad (70)$$

$$z = t + \frac{19}{7} \quad (71)$$

To find where this line meets the plane $3x + 8y - 4z = 7$, we find the value of t where the point described by the above line is also in the plane:

$$3 \left(-17t + \frac{41}{7} \right) + 56t - 4 \left(t + \frac{19}{7} \right) = 7 \quad (72)$$

$$-51t + 52t = 7 + \frac{76}{7} - \frac{123}{7} \quad (73)$$

$$t = \frac{49 + 76 - 123}{7} = \frac{2}{7} \quad (74)$$

- **Exercise 22** Prove that the distance D between a point \mathbf{x} and the plane $\mathbf{a} \cdot \mathbf{x} + d = 0$ is

$$D = \frac{|\mathbf{a} \cdot \mathbf{x} + d|}{|\mathbf{a}|} \quad (75)$$

- **Exercise 23** Using vector methods, find the equation of the straight line perpendicular to the line $3x + 4y - 2 = 0$ at the point $(6, -4)$.
- **Exercise 24** Find the (shortest) distance of the point $(3, 4, 5)$ from the plane through $(1, 2, 3)$, $(3, 2, 1)$ and $(6, 6, 1)$.
- **Exercise 25** Given a line described by the equation $ax + by + c = 0$, write down the unit vector perpendicular to this line.

2 Introductory Linear Algebra

2.1 Matrices

A Matrix is a rectangular array of numbers (scalars). The numbers in the array are called entries in the matrix.

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \\ -10 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ -1 & 3 & 5 \end{pmatrix}, \quad (1) \quad (76)$$

The size (or order) of the matrix is specified as $m \times n$, where m is the number of rows, and n is the number of columns.

Note that this means that n -dimensional (column) vectors are also $n \times 1$ matrices. Similarly, n -dimensional row vectors are also $1 \times n$ matrices.

2.2 Operations on Matrices

Matrix Addition

Matrices of the same size can be added. The summation is carried out as follows: if a is the (m, n) entry of matrix A , and b is the (m, n) entry of matrix B , then $a + b$ is the (m, n) entry of matrix $A + B$.

For example, if A and B are as defined below

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 6 & 5 & 2 \end{pmatrix} \quad B = \begin{pmatrix} 4 & 2 & 1 \\ 4 & -1 & 7 \end{pmatrix} \quad (77)$$

then

$$A + B = \begin{pmatrix} 5 & 4 & 4 \\ 10 & 4 & 9 \end{pmatrix} \quad (78)$$

Matrix Multiplication

If the elements of the $m \times r$ matrix A are a_{ij} , and the elements of the $r \times n$ matrix B are b_{ij} , then we define the elements of the $m \times n$ matrix AB as $[AB]_{ij} = \sum_{k=1}^r a_{ik}b_{kj}$.

For example, if the matrices are

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}, \quad (79)$$

Then

$$AB = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}, \tag{80}$$

Vectors can be regarded as matrices consisting of a single column. In this case, matrix multiplication follows the same rules as above, but simplifies somewhat. For example,

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_{11}b_1 + a_{12}b_2 + a_{13}b_3 \\ a_{21}b_1 + a_{22}b_2 + a_{23}b_3 \\ a_{31}b_1 + a_{32}b_2 + a_{33}b_3 \end{pmatrix} \tag{81}$$

Another example of matrix multiplication is given by

$$(a_1 \ a_2 \ a_3) \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = (a_1b_1 + a_2b_2 + a_3b_3) \tag{82}$$

If we identify a 1×1 matrix as a scalar, we see that the above equation is equivalent to $\mathbf{a} \cdot \mathbf{b}$, where $\mathbf{a} = (a_1, a_2, a_3)^T$ and $\mathbf{b} = (b_1, b_2, b_3)^T$.

- **Exercise 26** (Outer Product) For the matrices (vectors) $\mathbf{a} = (a_1, a_2, \dots, a_n)^T$ and $\mathbf{b} = (b_1, b_2, \dots, b_n)^T$ compute the (so called) vector outer product $\mathbf{a}\mathbf{b}^T$.

The scalar product in matrix notation

The above definitions means that we can write the scalar product in the alternative matrix notation form

$$\mathbf{a} \cdot \mathbf{b} \equiv \mathbf{a}^T \mathbf{b} \tag{83}$$

Multiplying a matrix by a scalar

A matrix A multiplied by a scalar k has elements ka_{ij} . We denote this as kA .

2.3 Some special Matrices

The Zero matrix

The zero matrix $\mathbf{0}$ is a matrix in which each element contains 0. Multiplication by the zero matrix produces the zero matrix. The zero matrix added to another matrix B reproduces B . Note that zero matrices need not be square (having an equal number of rows and columns).

The identity matrix

The identity matrix is defined as that matrix \mathbf{I} which, under matrix multiplication with any matrix A , leaves A unchanged. The identity matrix therefore has the form

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \tag{84}$$

that is, the diagonals of the matrix contain 1, and all other elements contain 0. Note that the identity matrix is necessarily square.

- **Exercise 27** If A and B are square matrices satisfying $A^2 = B$ show that A and B commute (*i.e.*, $AB = BA$). Hence find matrices A such that

$$A^2 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \tag{85}$$

- **Exercise 28** If $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 5 & 6 & 7 \\ 8 & 9 & 10 \end{pmatrix}$, calculate $A+B$, AB , BA and $B^T A$.
- **Exercise 29** If matrix X has order $p \times q$ and matrix Y has order $r \times s$, what is the order of the matrix XY ?
-

2.4 Linear Transformations of Vectors

The matrix multiplication ideas that we developed above are very useful in describing linear transformations of vectors.

Since matrix multiplication is distributive, one can understand (and, conversely, construct) transformations by the effect (or desired effect) on the unit axis vectors.

Example 8 What transformation does the matrix $A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ represent?

Solution Lets see what happens to the unit vectors \mathbf{i} and \mathbf{j} . \mathbf{i} is transformed to

$$A\mathbf{i} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad (86)$$

and \mathbf{j} is transformed to

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

(note that it is always true that the transformation of \mathbf{i} is the first column of the transformation matrix, and the transformation of \mathbf{j} is the second column of the matrix.) We can see therefore that A represents a rotation about 180 degrees.

Translation

If a set of (position) vectors \mathbf{x} is moved to a new position, simply by the addition of a constant vector, this is known as translation.

$$\mathbf{x}' = \mathbf{x} + \mathbf{a} \quad (87)$$

Scaling

Scaling a vector means changing the length of the vector, but not its direction. Since the direction in which a vector points is associated with the *ratio* of its components, scaling corresponds to simply multiplying all the components of the vector by an equal amount. That is

$$\mathbf{x}' = k\mathbf{x} \equiv S\mathbf{x} \quad (88)$$

where the scaling matrix $S = k\mathbf{I}$, k times the identity matrix.

Reflection

Reflection operators flip vectors about a certain line (in two dimensions) or plane (in three dimensions). For example, if we flip a vector $\mathbf{x} = (x, y)^T$ about the y -axis, this becomes

$$x' = -x' \quad y' = y' \quad (89)$$

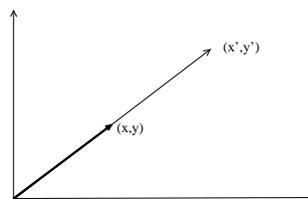


Figure 16: Scaling a vector.

In matrix notation, if we let $\mathbf{x}' = (x', y')^T$, this reflection is expressed as

$$\mathbf{x}' = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x} \quad (90)$$

Example 9 Find the matrix which represents a reflection about the plane $\mathbf{x} \cdot \mathbf{n} = 0$ where \mathbf{n} is a unit vector.

Solution Any vector \mathbf{x} can be written as

$$\mathbf{x} = \mathbf{r} + (\mathbf{x} \cdot \mathbf{n}) \mathbf{n} \quad (91)$$

where $(\mathbf{x} \cdot \mathbf{n})\mathbf{n}$ is the component of the vector perpendicular to the plane, and \mathbf{r} is the component parallel to the plane. Under the reflection, this becomes

$$\mathbf{x}' = \mathbf{r} - (\mathbf{x} \cdot \mathbf{n}) \mathbf{n} \quad (92)$$

or

$$\mathbf{x}' = \mathbf{x} - 2(\mathbf{x} \cdot \mathbf{n}) \mathbf{n} \quad (93)$$

The images of the unit vectors \mathbf{i} , \mathbf{j} and \mathbf{k} are the columns of the matrix representing this mapping. This is therefore

$$\begin{pmatrix} 1 - 2n_1^2 & -2n_1n_2 & -2n_1n_3 \\ -2n_1n_2 & 1 - 2n_2^2 & -2n_2n_3 \\ -2n_1n_3 & -2n_2n_3 & 1 - 2n_3^2 \end{pmatrix} \quad (94)$$

► **Exercise 30** Find the matrix which represents a reflection in two dimensions along the line $y = x \tan \alpha$.

Projection

We have already encountered examples of projections when looking at the scalar product. We saw there that a vector projected onto the plane specified by the unit vector \mathbf{n} can be found as follows.

A vector \mathbf{x} can be represented by a component parallel to \mathbf{r} and a component \mathbf{r}^\perp perpendicular to \mathbf{r} as

$$\mathbf{x} = (\mathbf{x} \cdot \mathbf{n}) \mathbf{n} + \mathbf{r}^\perp \quad (95)$$

Under the projection operator, the new vector contains only the component perpendicular to \mathbf{n}

$$\mathbf{x}' = \mathbf{r} - (\mathbf{x} \cdot \mathbf{n}) \mathbf{n} \quad (96)$$

By considering what happens to the unit vectors \mathbf{i} , \mathbf{j} and \mathbf{k} , the projection matrix is given by

$$\begin{pmatrix} 1 - n_1^2 & -n_1 n_2 & -n_1 n_3 \\ -n_1 n_2 & 1 - n_2^2 & -n_2 n_3 \\ -n_1 n_3 & -n_2 n_3 & 1 - n_3^2 \end{pmatrix} \quad (97)$$

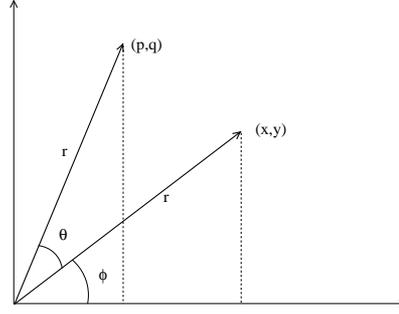


Figure 17: Rotating a vector through θ degrees.

Rotation

From trigonometry

$$x = r \cos \phi, \quad y = r \sin \phi \quad (98)$$

and

$$p = r \cos (\theta + \phi), \quad q = r \sin (\theta + \phi) \quad (99)$$

Using the identities for the addition of angles, the above can be written as

$$p = r \cos \theta \cos \phi - r \sin \theta \sin \phi \quad (100)$$

$$q = r \sin \theta \cos \phi + r \cos \theta \sin \phi \quad (101)$$

Using the definition of the coordinates x, y , this is

$$p = x \cos \theta - y \sin \theta \quad (102)$$

$$q = x \sin \theta + y \cos \theta \quad (103)$$

In matrix notation, this is

$$\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (104)$$

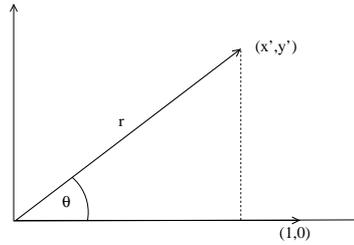
Or, if we write $\mathbf{x}' = (p, q)^T$ and $\mathbf{x} = (x, y)^T$, this transformation can be written succinctly as $\mathbf{x}' = R\mathbf{x}$ where R is the rotation matrix

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (105)$$

Rotations – the elegant way

Note that there is a more elegant way of deriving this result for the form of the matrix that rotates vectors through a certain angle.

If we assume that rotation of a vector can be accomplished by matrix multiplication of the vector, then, since matrix multiplication is distributive, we only need to work out how the axes unit vectors \mathbf{i} and \mathbf{j} transform.

Figure 18: Rotating a vector through θ degrees.

That is,

$$R\mathbf{x} = xR\mathbf{i} + yR\mathbf{j} \quad (106)$$

How can we find $R\mathbf{i}$? This situation is depicted in fig(18). The unit vector $(1, 0)^T$, under rotation by θ degrees gets transformed to the vector

$$R\mathbf{i} = \begin{pmatrix} r_{11} \\ r_{21} \end{pmatrix} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \quad (107)$$

Similarly, the unit vector \mathbf{j} transforms according to

$$R\mathbf{j} = \begin{pmatrix} r_{12} \\ r_{22} \end{pmatrix} = \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix} \quad (108)$$

From this, one can simply read off the values for the elements of R .

2.5 Compositions of Linear Transformations

All the above transformations can be represented by matrix multiplications or additions on vectors. We can apply a sequence of such transformations, for example, a rotation R , followed by a rescaling, S . This would have the effect of transforming a vector \mathbf{x} to a new vector \mathbf{x}' ,

$$\mathbf{x}' = S(R\mathbf{x}) \quad (109)$$

Since matrix multiplication is associative (because normal arithmetic is),

$$\mathbf{x}' = S(R\mathbf{x}) \equiv (SR)\mathbf{x} \quad (110)$$

so that we can equivalently, first calculate the matrix product SR , and subsequently apply this new matrix to the vector \mathbf{x} .

► **Exercise 31** Use the fact that a matrix representing a rotation through θ degrees multiplied by a rotation matrix through ϕ degrees is equivalent to a rotation matrix through $\theta + \phi$ to find expressions for $\sin(\theta + \phi)$ and $\cos(\theta + \phi)$.

► **Exercise 32** If

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 1 & -3 & 3 & -1 \end{pmatrix} \quad (111)$$

prove that $A^2 = \mathbf{I}$, where \mathbf{I} is the (4×4) unit matrix. Prove also that if

$$P = AM_1A \quad \text{and} \quad Q = AM_2A \quad (112)$$

where M_1 and M_2 are arbitrary diagonal matrices of order (4×4) , then $PQ = QP$.

- **Exercise 33** The following matrices represent linear transformations. Plot the transformations of the vectors $(1, 0)^T$ and $(0, 1)^T$, and hence classify each transformation.

$$\begin{pmatrix} 2 & 0 \\ 3 & 0 \end{pmatrix} \quad \begin{pmatrix} 2 & 3 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \quad \begin{pmatrix} 0 & 2 \\ 3 & 0 \end{pmatrix} \quad (113)$$

- **Exercise 34** Given $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 2 & 3 \\ -1 & 2 \\ 2 & 1 \end{pmatrix}$ and $C = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ find $(AB)C$ and $A(BC)$ and comment.

- **Exercise 35** Find the matrix which represents, in three dimensions, a rotation of 180° about the x -axis.

2.6 Determinants

We define (loosely) the determinant as (up to a sign change) the volume of the transformation of the matrix A . That is, we take a hypercube of unit volume and map each vertex under the transformation, and the volume of the resulting object is defined as the determinant of the transformation.

We shall explicitly evaluate the determinant for initially two and then three dimensions, from which we then state the generalisation to arbitrary dimensions. From our previous examination of matrices, we know that the

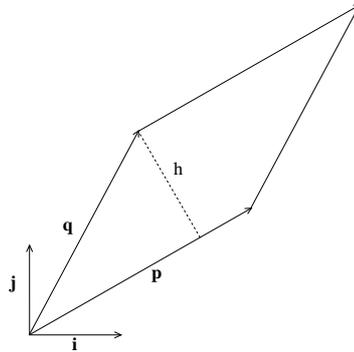


Figure 19: To find the area of the parallelogram formed by \mathbf{p} and \mathbf{q} , we need to know the height h of the equivalent area rectangle. The area is then given by $|\mathbf{q}|h$.

transformation can be wholly understood in terms of the effect that the transformation has on the unit vectors \mathbf{i} and \mathbf{j} . We write the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = (\mathbf{p} \quad \mathbf{q}) \quad (114)$$

where $\mathbf{p} = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ and $\mathbf{q} = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$. This means that the unit vector \mathbf{i} is transformed to \mathbf{p} , and \mathbf{j} is transformed to \mathbf{q} . The volume of this transformation is $|\mathbf{p}|h$ where h is the perpendicular distance from \mathbf{p} to \mathbf{q} , as shown in fig(19).

This area is given by $a_{11}a_{22} - a_{21}a_{12}$ (exercise), hence

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{21}a_{12} \quad (115)$$

Similarly, in three dimensions, the volume of the transformation is given by the triple product (not described in these notes) of the three vectors that form the columns of the matrix

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}) \quad (116)$$

You may be able to spot a pattern here in the form of the determinants. The determinant in the (3×3) case has the form

$$a_{11} \det \begin{pmatrix} a_{12} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \quad (117)$$

That is, the determinant of the (3×3) matrix A is given by the sum of terms $(-1)^{i+1}a_{1i} \det A_i$ where A_i is the (2×2) matrix formed from A by removing the i^{th} row and column.

Although we shall not prove this here, this form of the determinant generalises to any dimension. That is, we can define the determinant recursively as an expansion along the top row of determinants of reduced matrices.

For example,

$$\begin{aligned} & \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} \\ &= a_{11} \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} - a_{14} \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{33} & a_{33} \\ a_{41} & a_{43} & a_{43} \end{vmatrix} \end{aligned} \quad (118)$$

Each of the (3×3) determinants is evaluated using the formula equation (117).

Note that, with the above recursive definition, the determinant need not be a positive number (so, strictly this is not a volume). By taking the absolute value of the determinant, as defined by this recursion, we obtain the volume of the transformation.

Properties of the Determinant

The following are properties of the determinant. We shall not prove all these properties here, but most are intuitively clear, given our relation of the determinant to the volume of the transformation. Other properties can be proved from the mathematical recursive definition of the determinant given above.

- $\det A^T = \det A$

- Multiplying a single row or column of A by a scalar, k changes the determinant by a factor k . That is $\det B = k \det A$ where B is A but with one of the columns or rows multiplied by k .
- Interchanging two rows or two columns changes the sign of the determinant.
- If B is the matrix that results when a multiple of one row of A is added to another row or when a multiple of one column is added to another column, then $\det B = \det A$.
- If A is a square matrix with two proportional rows or columns then $\det A = 0$.
- If A and B are square matrices of the same size,

$$\det AB = \det A \det B \quad (119)$$

- The determinant of a triangular matrix (either all the elements above the diagonal or all the elements below the diagonal are zero) is equal to the product of its diagonal elements.

2.7 Matrix Inversion

There are several ways to understand matrix inversion. One of them is to build on the concepts of linear transformations that we have been developing. That is, if we applied the transformation A to a vector \mathbf{x} , what transformation do we need to apply to the vector $A\mathbf{x}$ to recover the original vector \mathbf{x} ? Mathematically, we are looking for a matrix, which we denote by A^{-1} such that

$$A^{-1}A\mathbf{x} = \mathbf{x} \quad (120)$$

since we want this to be true for any vector \mathbf{x} , what we require is

$$A^{-1}A = \mathbf{I} \quad (121)$$

Singular Matrices

Note that it is not always possible to find a matrix A^{-1} such that $A^{-1}A = \mathbf{I}$. In that case, we call the matrix A *singular*.

Inverses of matrix products

We will here show that $(AB)^{-1} = B^{-1}A^{-1}$.

Lets assume that this is true, which means that we need to show that

$$(AB)^{-1}AB = \mathbf{I} \quad (122)$$

Using the claimed form for the inverse, the left hand side is

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}\mathbf{I}B = B^{-1}B = \mathbf{I} \quad (123)$$

Equivalence of the right and left matrix inverse

Above we defined a matrix which, multiplied on the left of another matrix, gave the identity matrix. This is known as the left inverse. We will show that the corresponding right inverse is the same matrix as the left inverse.

Multiplying equation (121) on the left by A , and on the right by A^{-1} , we obtain

$$AA^{-1}AA^{-1} = AA^{-1} \quad (124)$$

bracketing the left hand side as

$$(AA^{-1})(AA^{-1}) = (AA^{-1}) \quad (125)$$

we see that

$$AA^{-1} = \mathbf{I} \quad (126)$$

so that A^{-1} acts as both a right and a left inverse – they are the same.

Understanding the inverse

Given that we understand vector multiplication by a matrix A as being equivalent to a sequence of rotations, scalings, and projections, taking an initial vector \mathbf{x} to a new vector \mathbf{x}' , the inverse matrix A^{-1} restores the vector \mathbf{x}' to \mathbf{x} .

Generally, this is possible. However, consider the case in which the transformation A involves a projection. This means that, given \mathbf{x}' , there is no unique solution for \mathbf{x} – it could lie anywhere along a whole line, parallel to the direction of the projection, see fig(20). Any transformation that in-

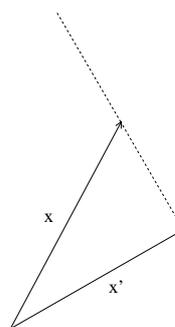


Figure 20: If we know the projection \mathbf{x}' of a vector, there is no unique solution to the question, “which vector was projected onto the vector \mathbf{x}' ?” Any vector with end point along the direction of the dashed line has the projection \mathbf{x}' .

volves a projection has volume zero, since (at least) one of the dimensions gets “flattened” to lie in the others. In other words, a matrix is invertible, provided that the volume of the transformation is non-zero.

Solving $A\mathbf{x} = \mathbf{b}$ algebraically

A set of simultaneous linear equations can always be written in the form

$$A\mathbf{x} = \mathbf{b} \quad (127)$$

Provided that the inverse of the matrix exists, then the solution is given by

$$\mathbf{x} = A^{-1}\mathbf{b} \quad (128)$$

Although this is, algebraically, the correct approach to solving this linear system of equations, in practice, we normally solve this problem using a numerical technique that explicitly avoids the need to do matrix inversion.

- **Exercise 36** Find the determinant of the matrix representing a projection on the plane specified by $\mathbf{x} \cdot \mathbf{n} = 0$ where \mathbf{n} is a unit vector.
- **Exercise 37** Let $A = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Find
- the set of all matrices X such that $X^2 = A$ (note that the notation X^2 means XX).
 - the set of all matrices Y such that $Y^2 = B$.
- **Exercise 38** Show that $(A + B)(A - B) = A^2 - B^2$ only if $AB = BA$ (this property of the matrices A and B is called commutability).
-

2.8 Computing the Matrix Inverse

Two dimensions

For a 2×2 matrix, it is straightforward to work out for a general matrix, the explicit form of the inverse. If the matrix whose inverse we wish to find is $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then the condition for the inverse is

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (129)$$

Multiplying out the left hand side, we obtain the four conditions

$$\begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (130)$$

It is readily verified that the solution to this set of four linear equations is given by

$$\begin{pmatrix} e & f \\ g & h \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = A^{-1} \quad (131)$$

The quantity $ad - bc$ is known as the determinant of the (2×2) matrix, and is denoted by $|A|$, or $\det A$. Although there is a clash with the notation for the length of a vector, it is always clear which is appropriate since the determinant of a vector does not make sense, as neither does the length of a (square) matrix.

General solution

There are many ways to compute the inverse of a matrix, but we will not go into this here since computers have made the inversion by hand of matrices redundant.

Note that, if one wants to solve only a linear system, although the solution can be obtained through matrix inversion, *this should never be applied!*. Often, in modern applications, one needs to solve huge dimensional linear systems of equations, and speed becomes an issue. These equations can be solved much more accurately and quickly using elimination techniques.

- **Exercise 39** Solve the following equations by matrix methods
- $4x - 3y + z = 11, 2x + y - 4z = -1, x + 2y - 2z = 1$
 - $x + 5y + 3z = 1, 5x + y - z = 2, x + 2y + z = 3$

- **Exercise 40** Solve

$$\begin{pmatrix} -1 & 3 & 4 \\ 1 & -4 & -5 \\ -2 & 3 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} \quad (132)$$

► **Exercise 41** Find the determinant of the matrix $\begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & 3 \\ 1 & -1 & 0 \end{pmatrix}$.

► **Exercise 42** Find the inverse of the matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & -1 \\ 3 & 8 & -4 \end{pmatrix} \quad (133)$$

and hence or otherwise solve

$$x + 2y + 3z = 14, 2x + 5y - z = 9, 3x + 8y - 4z = 7 \quad (134)$$

2.9 Eigenvalues and Eigenvectors

One of the most important concepts in linear algebra is that of eigenvalues and eigenvectors. In a sense, the eigenvectors of the matrix correspond to the natural coordinate system, in which the transformation can be most easily understood.

We call \mathbf{e} an eigenvector and λ the corresponding eigenvalue of the matrix A if

$$A\mathbf{e} = \lambda\mathbf{e} \quad (135)$$

Eigenvectors are usually scaled to have unit length, though not always.

Finding Eigenvalues

In general, for an $(n \times n)$ dimensional matrix, there are (possibly including repetitions) n eigenvalues, each with a corresponding eigenvector. We can reform equation (135) as

$$(A - \lambda\mathbf{I})\mathbf{e} = \mathbf{0} \quad (136)$$

This is simply a linear equation, and we can use our knowledge of linear algebra to solve this. We can write equation (136) as $B\mathbf{e} = \mathbf{0}$, where $B = A - \lambda\mathbf{I}$. If B has an inverse, then the solution is $\mathbf{e} = B^{-1}\mathbf{0} = \mathbf{0}$, since any matrix multiplying the zero vector gives zero. For any non-trivial solution to the problem $B\mathbf{e} = \mathbf{0}$, we therefore need B to be non-invertible. This is equivalent to the condition that the B has zero determinant.

In other words, λ is an eigenvalue of A if

$$\det(A - \lambda\mathbf{I}) = 0 \quad (137)$$

This is known as the characteristic equation. This determinant equation will be a polynomial of degree n (the dimension of A), and the resulting equation is known as the characteristic polynomial.

Once we have found an eigenvalue, the corresponding eigenvector can be found by substituting this value for λ in equation (135) and solving for \mathbf{e} .

Example 10 Find the eigenvalues and eigenvectors of $\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$

Solution The characteristic equation is

$$\begin{vmatrix} 1 - \lambda & 3 \\ 4 & 2 - \lambda \end{vmatrix} = 0 \quad (138)$$

Evaluating the determinant, this is

$$(1 - \lambda)(2 - \lambda) - 3 \times 4 = 0 \quad (139)$$

Factorising this quadratic, we have

$$(\lambda - 5)(\lambda + 2) = 0 \quad (140)$$

so that the solutions are $\lambda = 5$ and $\lambda = -2$. To find the eigenvectors, we consider each eigenvalue separately.

$\lambda = 5$

$$\begin{pmatrix} -4 & 3 \\ 4 & -3 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (141)$$

This is

$$-4e_1 + 3e_2 = 0, 4e_1 - 3e_2 = 0 \quad (142)$$

(Note that the second equation is simply the negative of the first. Such redundancies will always occur in the solution of these eigenvector equations). This means that the eigenvector points along $\begin{pmatrix} 1 \\ \frac{4}{3} \end{pmatrix}$ so that the unit eigenvector is (by normalising this vector)

$$\begin{pmatrix} \frac{3}{5} \\ \frac{4}{5} \end{pmatrix} \quad (143)$$

Similarly, we obtain for the other eigenvalue $\lambda = -2$

$$\begin{pmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \quad (144)$$

► **Exercise 43** Find the eigenvalues of a two dimensional rotation matrix through θ degrees. Comment on your result.

► **Exercise 44** Find a general expression for the eigenvalues of a general 2×2 matrix. Relate the product of the eigenvalues to the determinant of the matrix.

► **Exercise 45** Find the characteristic polynomial of the matrix

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 3 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix} \quad (145)$$

and prove that the matrices

$$B = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \text{and} \quad C = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 1 \end{pmatrix} \quad (146)$$

have the same characteristic polynomials as A .

► **Exercise 46** Show that if A has eigenvalues $\lambda_1 \dots \lambda_n$ then A^m (where m is a positive integer) has eigenvalues $\lambda_1^m \dots \lambda_n^m$

► **Exercise 47** Show that if A and B are of order n and A is a non-singular matrix then $A^{-1}BA$ and B have the same eigenvalues.

Example 11 For example, lets find the eigenvalues and eigenvectors of the matrix:

$$\begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix} \quad (147)$$

First we need to solve the eigenvalue problem

$$\begin{vmatrix} -\lambda & 0 & -2 \\ 1 & 2-\lambda & 1 \\ 1 & 0 & 3-\lambda \end{vmatrix} = 0 \quad (148)$$

The characteristic polynomial is

$$\lambda^3 - 5\lambda^2 + 8\lambda - 4 = 0 \quad (149)$$

We need to solve this equation. One good way at guessing a solution is to use the fact that, (as is always true), the product of the roots of a polynomial is equal to the constant in the polynomial. If we guess that the roots are integers then, since their product must be equal to 4, the integer roots could be 1, 2, or 4. Substituting in $\lambda = 1$ in the polynomial, we see this is a solution. Similarly, $\lambda = 2$ is a solution. Since the product of the roots must be equal to four, the only possibility for the other root is that it must also be equal to 2. That is, there is a repeated eigenvalue, namely $\lambda = 2$.

Lets try to find the eigenvectors for $\lambda = 1$:

$$\begin{pmatrix} -1 & 0 & -2 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (150)$$

From the top line, we have $e_1 = -2e_2$. From the second line, $e_3 = -(e_1 + e_2)$. If we set $e_1 = s$, then $e_2 = -s/2$ and $e_3 = -s/2$. This means that the eigenspace is spanned by the basis

$$\begin{pmatrix} 1 \\ -1/2 \\ -1/2 \end{pmatrix} \quad (151)$$

For $\lambda = 2$, the situation is a little more complicated

$$\begin{pmatrix} -2 & 0 & -2 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (152)$$

As we can see, if we set $e_1 = s$, then we have the condition $e_3 = -s$. However, e_2 is left free. This means that any vector \mathbf{e} in the space

$$s \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad (153)$$

is an eigenvector of A with eigenvalue 2. We can characterise this eigenspace by a set of basis vectors that span the space, namely

$$\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \text{ and } \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad (154)$$

Relating the eigenvalues to the determinant

The determinant of an $n \times n$ matrix A is the product of its eigenvalues,

$$\det A = \prod_{i=1}^n \lambda_i \quad (155)$$

- **Exercise 48** By considering the definition of the characteristic polynomial, show that the determinant of a matrix is equal to the product its eigenvalues.

The orthogonality of eigenvectors of a Symmetric Matrix

Consider two distinct eigenvectors of a symmetric matrix A .

$$A\mathbf{x}_1 = \lambda_1\mathbf{x}_1 \quad A\mathbf{x}_2 = \lambda_2\mathbf{x}_2 \quad (156)$$

Multiplying (say) the left equation by \mathbf{x}_2^T , we obtain

$$\mathbf{x}_2^T A\mathbf{x}_1 = \lambda_1 \mathbf{x}_2^T \mathbf{x}_1 \quad (157)$$

This is a scalar equation, and is therefore equal to its transpose,

$$\mathbf{x}_1^T A^T \mathbf{x}_2 = \lambda_1 \mathbf{x}_1^T \mathbf{x}_2 \quad (158)$$

Using the symmetry of A , $A = A^T$, and $A\mathbf{x}_2 = \lambda_2\mathbf{x}_2$, this becomes

$$\lambda_2 \mathbf{x}_1^T \mathbf{x}_2 = \lambda_1 \mathbf{x}_1^T \mathbf{x}_2 \quad (159)$$

Since we assumed that $\lambda_1 \neq \lambda_2$, the only solution to this is that both sides are zero, namely $\mathbf{x}_1^T \mathbf{x}_2 = 0$. In other words, the eigenvectors of a symmetric matrix form a mutually orthogonal set.

Outer Product Representation of a symmetric matrix

$$A = \sum_{i=1}^n \lambda_i \mathbf{e}_i \mathbf{e}_i^T \quad (160)$$

where λ_i is the eigenvalue of eigenvector \mathbf{e}_i .

- **Exercise 49** Explain why any symmetric matrix has such an outer product representation.
-

Answer 1 $AC = a + 3b$, $DB = 3b - a$, $BC = 2a + 2b$, $CD = -a - 5b$

Answer 2 $(1, -1)$, $(1, -1, 1)$

Answer 3 $(2, -9, -7)$, $(2, 9, 0)$, $(-6 + 2\alpha, -11 + \alpha, -7)$

Answer 4 $\alpha = 5/13$, $\beta = 6/13$, $\gamma = 1/13$

Answer 5 $\alpha = 3 + -2/\sqrt{3}$

Answer 6 The boat's position at time t is

$$\mathbf{p} = t \left(4 \begin{pmatrix} \sin 40 \\ \cos 40 \end{pmatrix} + \mathbf{c} \right)$$

where \mathbf{c} is the velocity of the water. After one hour,

$$4 \begin{pmatrix} \sin 40 + c_1 \\ \cos 40 + c_2 \end{pmatrix} = 6.12 \begin{pmatrix} \sin 60 \\ \cos 60 \end{pmatrix}$$

This gives $c_1 = 2.7289$, $c_2 = -0.0042$.

Answer 9 $3, \sqrt{8}, \sqrt{13} \cos \theta = (2/(3\sqrt{8}), 7/(3\sqrt{13}), 6/\sqrt{104})$

Answer 10 120^0

Answer 12 $\cos \theta = \sqrt{2/3} = 35^0$

Answer 13 See fig(21)

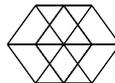


Figure 21:

Answer 14 The answer is $N2^{N-1}$. There are 2^N vertices, and each vertex is connected to N others. However, we double count the number of edges in this way, since each edge belongs to two vertices.

Answer 15 $(0, 0), (8/13, -12/13), (19/13, 0), (95/13), (16/89, 12/89, 32/89)$

Answer 16 $(6, 2), (-21/13, -14/13), (20/13, 1, -4/13), (73/89, -12/89, -32/89)$

Answer 17 $-36, 13$

Answer 18 First find a formula for calculating the distance (shortest) from a point to a line. We can then find that point on the line with shortest distance. If \mathbf{u} is a unit vector along the line, then any point in space can be written as

$$\mathbf{p} = \mathbf{a} + \lambda \mathbf{u} + \mathbf{v} \tag{161}$$

where \mathbf{a} lies on the line, and \mathbf{v} is perpendicular to the line. Taking the scalar product with \mathbf{u} gives the projection onto the line, $\mathbf{a} + \lambda \mathbf{u}$ as

$$\mathbf{q} = \mathbf{a} + \lambda \mathbf{u} = \mathbf{a} + (\mathbf{p} - \mathbf{a}) \cdot \mathbf{u} \mathbf{u} \tag{162}$$

So that the shortest square distance is given

$$(\mathbf{q} - \mathbf{p})^2 = (\mathbf{p} - \mathbf{a})^2 - ((\mathbf{p} - \mathbf{a}) \cdot \mathbf{u})^2 \tag{163}$$

If we assume that \mathbf{p} is itself parameterised by $\mathbf{p} = \mathbf{c} + s\mathbf{d}$, the distance is

$$s^2 (\mathbf{d}^2 - (\mathbf{d} \cdot \mathbf{u})^2) + 2s (\mathbf{c} \cdot \mathbf{d} - \mathbf{d} \cdot \mathbf{a} - \mathbf{c} \cdot \mathbf{u} \mathbf{d} \cdot \mathbf{u} + \mathbf{d} \cdot \mathbf{u} \mathbf{a} \cdot \mathbf{u}) + (\mathbf{c} - \mathbf{a})^2 - ((\mathbf{c} - \mathbf{a}) \cdot \mathbf{u})^2$$

or

$$s^2 (\mathbf{d}^2 - (\mathbf{d} \cdot \mathbf{u})^2) + 2s (\mathbf{c} - \mathbf{a}) \cdot (\mathbf{d} - \mathbf{d} \cdot \mathbf{u} \mathbf{u}) + (\mathbf{c} - \mathbf{a})^2 - ((\mathbf{c} - \mathbf{a}) \cdot \mathbf{u})^2 \tag{164}$$

This is of the form $as^2 + 2bs + c$ which has a minimum value of $c - b^2/a$. This is then

$$(\mathbf{c} - \mathbf{a})^2 - ((\mathbf{c} - \mathbf{a}) \cdot \mathbf{u})^2 - \frac{((\mathbf{c} - \mathbf{a}) \cdot (\mathbf{d} - \mathbf{d} \cdot \mathbf{u} \mathbf{u}))^2}{\mathbf{d}^2 - (\mathbf{d} \cdot \mathbf{u})^2} \tag{165}$$

In this case, $\mathbf{a} = (13 - 2)^T$, $\mathbf{u} = (425)/\sqrt{45}$, $\mathbf{c} = (21 - 2)$ and $\mathbf{d} = (-223)$. I get the squared distance as 2.4845 or dist= 1.5762.

Answer 19 $4x - 5y = 1$

Answer 21 $x_3 b_3 + \beta = 0$ $p = (0, 0, -\beta/b_3)$, distance = $\mathbf{p} \cdot \hat{\mathbf{a}} + \alpha = -\frac{\beta a_3}{b_3 \sqrt{a_1^2 + a_2^2 + a_3^2}}$.

Answer 23 $3y - 4x + 36 = 0$

Answer 24 $10/\sqrt{41} = 1.56$

Answer 27 By using $BA = AB$, we get that A must be of the form $A = \begin{pmatrix} a & b & 0 \\ 0 & a & 0 \\ 2b & h & a \end{pmatrix}$.

Using $A^2 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$, we get $a^2 = 1, 2ab = 1, b^2 + ah = 0$. Solving gives

$$A = (+/-) \begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1 & 0 \\ 0 & -1/4 & 1 \end{pmatrix}$$

Answer 28 All rubbish except $AB = \begin{pmatrix} 21 & 24 & 27 \\ 47 & 54 & 61 \end{pmatrix}$, $B^T A = \begin{pmatrix} 29 & 42 \\ 33 & 48 \\ 37 & 54 \end{pmatrix}$.

Answer 29 $p \times r$

Answer 30 $\begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix}$

Answer 33 None!

Answer 34 $\begin{pmatrix} 36 \\ 27 \end{pmatrix}$ for both – demonstrates associativity of matrix multiplication.

Answer 35 $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

Answer 36 0

Answer 37 (a) $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix}$ (b) no matrices.

Answer 39 (a) $x = 3, y = 1, z = 2$. (b) inconsistent

Answer 40 $(3, -1, 0)^T$

Answer 41 6

Answer 42

$$\begin{pmatrix} -12 & 32 & -17 \\ 5 & -13 & 7 \\ 1 & -2 & 1 \end{pmatrix} \quad (166)$$

Answer 43 Unless the rotation corresponds to 0, 90, 180, or 270 degrees, then the eigenvalues are complex.

Answer 45 $\lambda^3 - 6\lambda^2 - 3\lambda + 18 = 0$

Answer 48 The constant term in the characteristic polynomial equals the matrix determinant. This is also the product of the roots of the polynomial.