Certainly one such possibility is

$$r^{\frac{1}{n}}\angle\frac{\theta}{n},$$

by virtue of the paragraph dealing with positive whole number powers.

But the general expression for z is given by

$$z = r\angle(\theta + k360^\circ),$$

where k may be any integer; and this suggests other possibilities for $z^{\frac{1}{n}}$, namely

$$r^{\frac{1}{n}}\angle\frac{\theta + k360^\circ}{n}.$$

However, this set of n-th roots is not an infinite set because the roots which are given by $k = 0, 1, 2, 3............n - 1$ are also given by $k = n, n + 1, n + 2, n + 3, ..., 2n - 1, 2n,$ $2n + 1, 2n + 2, 2n + 3, ...$ and so on, respectively.

We conclude that there are precisely n n-th roots given by $k = 0, 1, 2, 3........., n - 1$.

EXAMPLE

Determine the cube roots (i.e. 3rd roots) of the complex number $j8$.

Solution

We first write

$$j8 = 8\angle(90^\circ + k360^\circ).$$

Hence,

$$(j8)^{\frac{1}{3}} = 8^{\frac{1}{3}}\angle\frac{(90^\circ + k360^\circ)}{3},$$

where $k = 0, 1, 2$

The three distinct cube roots are therefore

$$2\angle 30^\circ, 2\angle 150^\circ \quad \text{and} \quad 2\angle 270^\circ = 2\angle(-90^\circ).$$

They all have the same modulus of 2 but their arguments are spaced around the Argand Diagram at regular intervals of $\frac{360^\circ}{3} = 120^\circ$.

Notes:

(i) In general, the n-th roots of a complex number will all have the same modulus, but their arguments will be spaced at regular intervals of $\frac{360^\circ}{n}$.

(ii) Assuming that $-180^\circ < \theta \leq 180^\circ$; that is, assuming that the polar form of z uses the principal value of the argument, then the particular n-th root of z which is given by $k = 0$ is called the **"principal n-th root"**.

(iii) If $\frac{m}{n}$ is a fraction in its lowest terms, we define

$$z^{\frac{m}{n}}$$

to be either $\left(z^{\frac{1}{n}}\right)^m$ or $(z^m)^{\frac{1}{n}}$ both of which turn out to give the same set of n distinct results.

The discussion, so far, on powers of complex numbers leads us to the following statement:

DE MOIVRE'S THEOREM

If $z = r\angle\theta$, then, for any rational number n, **one value** of z^n is $r^n\angle n\theta$.

6.4.4 EXERCISES

1. Determine the following in the form $a + jb$, expressing a and b in decimals correct to four significant figures:

   (a)
   $$(1 + j\sqrt{3})^{10};$$

   (b)
   $$(2 - j5)^{-4}.$$

2. Determine the fourth roots of $j81$ in exponential form $re^{j\theta}$ where $r > 0$ and $-\pi < \theta \leq \pi$.
3. Determine the fifth roots of the complex number $-4 + j4$ in the form $a + jb$ expressing a and b in decimals, where appropriate, correct to two places. State also which root is the principal root.

4. Determine all the values of

$$(3 + j4)^{\frac{3}{2}}$$

in polar form.

6.4.5 ANSWERS TO EXERCISES

1. (a)

$$(1 + j\sqrt{3})^{10} = -512.0 - j886.8;$$

(b)

$$(2 - j5)^{-4} = 5.796 - j1.188$$

2. The fourth roots are

$$3e^{-\frac{\pi}{8}}, \quad 3e^{\frac{3\pi}{8}}, \quad 3e^{\frac{7\pi}{8}}, \quad 3e^{-\frac{5\pi}{8}}.$$

3. The fifth roots are

$$1.26 + j0.64, \quad -0.22 + j1.40, \quad -1.40 + j0.22, \quad -0.64 - j1.26, \quad 1 - j.$$

The principal root is $1.26 + j0.64$.

4. There are two values, namely

$$11.18\angle 79.695^\circ \quad \text{and} \quad 11.18\angle(-100.305^\circ).$$

"JUST THE MATHS"

UNIT NUMBER

6.5

COMPLEX NUMBERS 5
(Applications to trigonometric identities)

by

A.J.Hobson

UNIT 6.5 - COMPLEX NUMBERS 5

APPLICATIONS TO TRIGONOMETRIC IDENTITIES

6.5.1 INTRODUCTION

It will be useful for the purposes of this section to restate the result known as **"Pascal's Triangle"** previously discussed in Unit 2.2.

If n is a positive whole number, the diagram

$$\begin{array}{ccccccccc} & & & 1 & & 1 & & & \\ & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ 1 & & 4 & & 6 & & 4 & & 1 \end{array}$$

provides the coefficients in the expansion of $(A+B)^n$ which contains the sequence of terms

$$A^n,\ A^{n-1}B,\ A^{n-2}B^2,\ A^{n-3}B^3, ..., B^n.$$

6.5.2 EXPRESSIONS FOR $\cos n\theta$ AND $\sin n\theta$ IN TERMS OF $\cos\theta$ AND $\sin\theta$.

From De Moivre's Theorem

$$(\cos\theta + j\sin\theta)^n \equiv \cos n\theta + j\sin n\theta,$$

from which we may deduce that, in the expansion of the left-hand-side, using Pascal's Triangle, the real part will coincide with $\cos n\theta$ and the imaginary part will coincide with $\sin n\theta$.

EXAMPLE

$$(\cos\theta + j\sin\theta)^3 \equiv \cos^3\theta + 3\cos^2\theta.(j\sin\theta) + 3\cos\theta.(j\sin\theta)^2 + (j\sin\theta)^3.$$

That is,

$$\cos 3\theta \equiv \cos^3\theta - 3\cos\theta.\sin^2\theta \quad \text{or} \quad 4\cos^3\theta - 3\cos\theta,$$

using $\sin^2\theta \equiv 1 - \cos^2\theta$;

and

$$\sin 3\theta \equiv 3\cos^2\theta.\sin\theta - \sin^3\theta \quad \text{or} \quad 3\sin\theta - 4\sin^3\theta,$$

using $\cos^2\theta \equiv 1 - \sin^2\theta$.

6.5.3 EXPRESSIONS FOR $\cos^n\theta$ AND $\sin^n\theta$ IN TERMS OF SINES AND COSINES OF WHOLE MULTIPLES OF θ.

The technique described here is particularly useful in calculus problems when we are required to integrate an integer power of a sine function or a cosine function. It does stand, however, as a self-contained application to trigonometry of complex numbers.

Suppose

$$z \equiv \cos\theta + j\sin\theta \quad - \quad (1)$$

Then, by De Moivre's Theorem, or by direct manipulation,

$$\frac{1}{z} \equiv \cos\theta - j\sin\theta \quad - \quad (2).$$

Adding (1) and (2) together, then subtracting (2) from (1), we obtain

$z + \frac{1}{z} \equiv 2\cos\theta$	$z - \frac{1}{z} \equiv j2\sin\theta$

Also, by De Moivre's Theorem,

$$z^n \equiv \cos n\theta + j\sin n\theta \quad - \quad (3)$$

and

$$\frac{1}{z^n} \equiv \cos n\theta - j\sin n\theta \quad - \quad (4).$$

Adding (3) and (4) together, then subtracting (4) from (3), we obtain

$z^n + \frac{1}{z^n} \equiv 2\cos n\theta$	$z^n - \frac{1}{z^n} \equiv j2\sin n\theta$

We are now in a position to discuss some examples on finding trigonometric identities for whole number powers of $\sin\theta$ or $\cos\theta$.

EXAMPLES

1. Determine an identity for $\sin^3\theta$.

   Solution

   We use the result

$$j^3 2^3 \sin^3\theta \equiv \left(z - \frac{1}{z}\right)^3,$$

   where $z \equiv \cos\theta + j\sin\theta$.

   That is,

$$-j8\sin^3\theta \equiv z^3 - 3z^2.\frac{1}{z} + 3z.\left(\frac{1}{z}\right)^2 - \frac{1}{z^3}$$

   or, after cancelling common factors,

$$-j8\sin^3\theta \equiv z^3 - 3z + \frac{3}{z} - \frac{1}{z^3} \equiv \left(z^3 - \frac{1}{z^3}\right) - 3\left(z - \frac{1}{z}\right),$$

   which gives

$$-j8\sin^3\theta \equiv j2\sin 3\theta - j6\sin\theta.$$

   Hence,

$$\sin^3\theta \equiv \frac{1}{4}\left(3\sin\theta - \sin 3\theta\right).$$

2. Determine an identity for $\cos^4\theta$.

   Solution

   We use the result

$$2^4\cos^4\theta \equiv \left(z + \frac{1}{z}\right)^4,$$

   where $z \equiv \cos\theta + j\sin\theta$.

   That is,

$$16\cos^4\theta \equiv z^4 + 4z^3.\frac{1}{z} + 6z^2.\left(\frac{1}{z}\right)^2 + 4z.\left(\frac{1}{z}\right)^3 + \left(\frac{1}{z}\right)^4$$

   or, after cancelling common factors,

$$16\cos^4\theta \equiv z^4 + 4z^2 + 6 + \frac{4}{z^2} + \frac{1}{z^4} \equiv z^4 + \frac{1}{z^4} + 4\left(z^2 + \frac{1}{z^2}\right) + 6,$$

   which gives

$$16\cos^4\theta \equiv 2\cos 4\theta + 8\cos 2\theta + 6.$$

   Hence,

$$\cos^4\theta \equiv \frac{1}{8}\left(\cos 4\theta + 4\cos 2\theta + 3\right)$$

6.5.4 EXERCISES

1. Use a complex number method to determine identities for $\cos 4\theta$ and $\sin 4\theta$ in terms of $\sin\theta$ and $\cos\theta$.
2. Use a complex number method to determine an identity for $\sin^5\theta$ in terms of sines of whole multiples of θ.
3. Use a complex number method to deternine an identity for $\cos^6\theta$ in terms of cosines of whole multiples of θ.

6.5.5 ANSWERS TO EXERCISES

1.

$$\cos 4\theta \equiv \cos^4\theta - 6\cos^2\theta.\sin^2\theta$$

and

$$\sin 4\theta \equiv 4\cos^3\theta.\sin\theta - 4\cos\theta.\sin^3\theta.$$

2.

$$\sin^5\theta \equiv \frac{1}{16}(\sin 5\theta - 5\sin 3\theta + 10\sin\theta].$$

3.

$$\cos^6\theta \equiv \frac{1}{32}(\cos 6\theta + 6\cos 4\theta + 15\cos 2\theta + 10).$$

"JUST THE MATHS"

UNIT NUMBER

6.6

COMPLEX NUMBERS 6
(Complex loci)

by

A.J.Hobson

UNIT 6.6 - COMPLEX NUMBERS 6

COMPLEX LOCI

6.6.1 INTRODUCTION

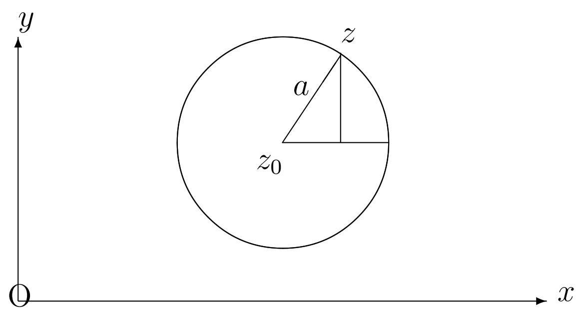
In Unit 6.2, it was mentioned that the directed line segment joining the point representing a complex number z_1 to the point representing a complex number z_2 is of length equal to $|z_2 - z_1|$ and is inclined to the positive direction of the real axis at an angle equal to $\arg(z_2 - z_1)$.

This observation now has significance when discussing variable complex numbers which are constrained to move along a certain path (or **"locus"**) in the Argand Diagram. For many practical applications, such paths (or **"loci"**) will normally be either straight lines or circles and two standard types of example appear in what follows.

In both types, we shall assume that $z = x + jy$ denotes a **variable** complex number (represented by the point (x, y) in the Argand Diagram), while $z_0 = x_0 + jy_0$ denotes a **fixed** complex number (represented by the point (x_0, y_0) in the Argand Diagram).

6.6.2 THE CIRCLE

Suppose that the moving point representing z moves on a circle, with radius a, whose centre is at the fixed point representing z_0.



Then the distance between these two points will always be equal to a. In other words,

$$|z - z_0| = a$$

and this is the standard equation of the circle in terms of complex numbers.

Note:
By substituting $z = x + jy$ and $z_0 = x_0 + jy_0$ in the above equation, we may obtain the equivalent equation in terms of cartesian co-ordinates, namely,

$$|(x - x_0) + j(y - y_0)| = a.$$

That is,

$$(x - x_0)^2 + (y - y_0)^2 = a^2.$$

ILLUSTRATION

The equation

$$|z - 3 + j4| = 7$$

represents a circle, with radius 7, whose centre is the point representing the complex number $3 - j4$.

In cartesian co-ordinates, it is the circle with equation

$$(x - 3)^2 + (y + 4)^2 = 49.$$

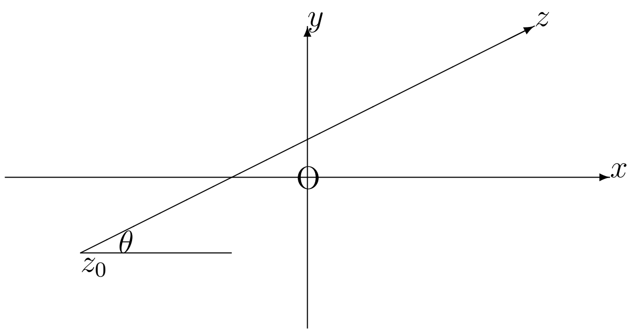
6.6.3 THE HALF-STRAIGHT-LINE

Suppose now that the **"directed"** straight line segment described **from** the fixed point representing z_0 **to** the moving point representing z is inclined at an angle θ to the positive direction of the real axis.

Then,

$$\arg(z - z_0) = \theta$$

and this equation is satisfied by **all** of the values of z for which the inclination of the directed line segment is genuinely θ and **not** $180^\circ - \theta$. The latter angle would correspond to points on the other half of the straight line joining the two points.



Note:
If we substitute $z = x + jy$ and $z_0 = x_0 + jy_0$, we obtain

$$\arg([x - x_0] + j[y - y_0]) = \theta.$$

That is,

$$\tan^{-1}\frac{y - y_0}{x - x_0} = \theta$$

or

$$y - y_0 = \tan\theta(x - x_0),$$

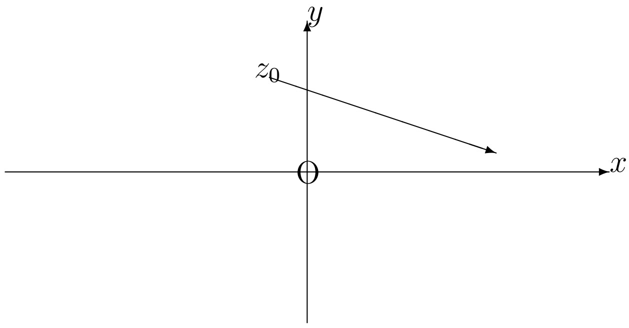
which is certainly the equation of a straight line with gradient $\tan\theta$ passing through the point (x_0, y_0); but it represents only that half of the straight line for which $x - x_0$ and $y - y_0$ correspond, in sign as well as value, to the real and imaginary parts of a complex number whose argument is genuinely θ and not $180^\circ - \theta$.

ILLUSTRATION

The equation

$$\arg(z + 1 - j5) = -\frac{\pi}{6}$$

represents the half-straight-line described from the point representing $z_0 = -1 + j5$ to the point representing $z = x + jy$ and inclined to the positive direction of the real axis at an angle of $-\frac{\pi}{6}$.



In terms of cartesian co-ordinates,

$$\arg([x + 1] + j[y - 5]) = -\frac{\pi}{6},$$

in which it must be true that $x + 1 > 0$ and $y - 5 < 0$ in order that the argument of $[x + 1] + j[y - 5]$ may be a negative acute angle.

We thus have the half-straight-line with equation

$$y - 5 = \tan\left(-\frac{\pi}{6}\right)(x + 1) = -\frac{1}{\sqrt{3}}(x + 1)$$

which lies to the right of, and below the point $(-1, 5)$.

6.6.4 MORE GENERAL LOCI

Certain types of locus problem may be encountered which cannot be identified with either of the two standard types discussed above. The secret, in such problems is to substitute $z = x + jy$ in order to obtain the cartesian equation of the locus. We have already seen that this method is applicable to the two standard types anyway.

ILLUSTRATIONS

1. The equation

$$\left|\frac{z-1}{z+2}\right| = 3$$

may be written

$$|z-1| = 3|z+2|.$$

That is,

$$(x-1)^2 + y^2 = 3[(x+2)^2 + y^2],$$

which simplifies to

$$2x^2 + 2y^2 + 14x + 13 = 0$$

or

$$\left(x + \frac{7}{2}\right)^2 + y^2 = \frac{23}{4},$$

representing a circle with centre $\left(-\frac{7}{2}, 0\right)$ and radius $\sqrt{\frac{23}{4}}$.

2. The equation

$$\arg\left(\frac{z-3}{z}\right) = \frac{\pi}{4}$$

may be written

$$\arg(z-3) - \arg z = \frac{\pi}{4}.$$

That is,

$$\arg([x-3]+jy)-\arg(x+jy)=\frac{\pi}{4}.$$

Taking tangents of both sides and using the trigonometric identity for $\tan(A-B)$, we obtain

$$\frac{\frac{y}{x-3}-\frac{y}{x}}{1+\frac{y}{x-3}\frac{y}{x}}=1.$$

On simplification, the equation becomes

$$x^2+y^2-3x-3y=0$$

or

$$\left(x-\frac{3}{2}\right)^2+\left(y-\frac{3}{2}\right)^2=\frac{9}{2},$$

the equation of a circle with centre $\left(\frac{3}{2},\frac{3}{2}\right)$ and radius $\frac{3}{\sqrt{2}}$.
However, we observe that the original complex number,

$$\frac{z-3}{z},$$

cannot have an argument of $\frac{\pi}{4}$ unless its real and imaginary parts are **both** positive. In fact,

$$\frac{z-3}{z}=\frac{(x-3)+jy}{x+jy}.\frac{x-jy}{x-jy}=\frac{x(x-3)+y^2+j3}{x^2+y^2}$$

which requires, therefore, that

$$x(x-3)+y^2>0.$$

That is,

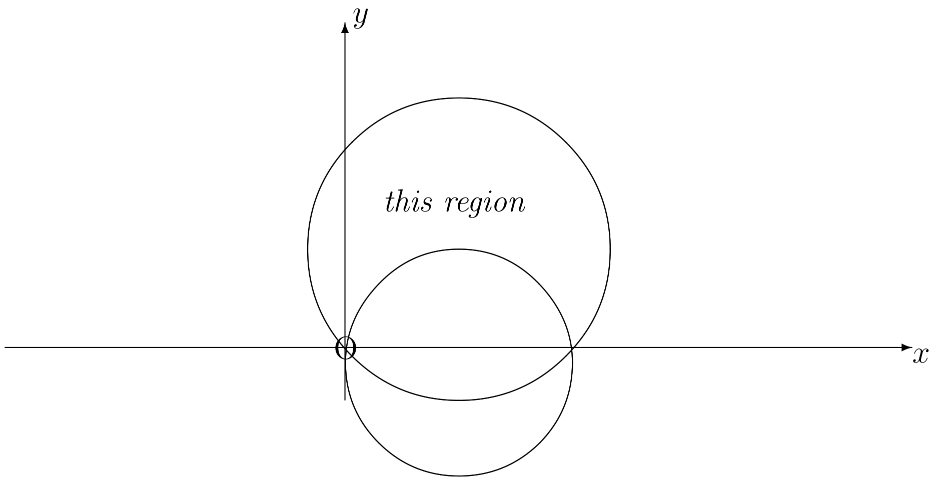
$$x^2+y^2-3x>0$$

or

$$\left(x-\frac{3}{2}\right)^2+y^2>\frac{9}{4}.$$

Conclusion

The locus is that part of the circle with centre $\left(\frac{3}{2},\frac{3}{2}\right)$ and radius $\frac{3}{\sqrt{2}}$ which lies **outside** the circle with centre $\left(\frac{3}{2},0\right)$ and radius $\frac{3}{2}$.



6.6.5 EXERCISES

1. Identify the loci whose equations are

   (a)
   $$|z-3|=4;$$

   (b)
   $$|z-4+j7|=2.$$

2. Identify the loci whose equations are

   (a)
   $$\arg(z+1)=\frac{\pi}{3};$$

   (b)
   $$\arg(z-2-j3)=\frac{3\pi}{2}.$$

3. Identify the loci whose equations are

   (a)
   $$\left|\frac{z+j2}{z-j3}\right|=1;$$

   (b)
   $$\arg\left(\frac{z+j}{z-1}\right)=-\frac{\pi}{4}.$$

6.6.6 ANSWERS TO EXERCISES

1. (a) A circle with centre $(3, 0)$ and radius 4;

   (b) A circle with centre $4, -7)$ and radius 2.

2. (a) A half-straight-line to the right of, and above the point $(-1, 0)$ inclined at an angle of $\frac{\pi}{3}$ to the positive direction of the real axis;

   (b) A half-straight-line below the point $(2, 3)$ and perpendicular to the real axis.

3. (a) The straight line $y = \frac{1}{2}$;

   (b) That part of the circle $x^2 + y^2 = 1$ which lies outside the circle with centre $\left(\frac{1}{2}, -\frac{1}{2}\right)$ and radius $\frac{1}{\sqrt{2}}$ **and** above the straight line whose equation is $y = x - 1$.

Note:
Examples like No. 3(b) are often quite difficult and will not normally be included in the more elementary first year courses in mathematics.

"JUST THE MATHS"

UNIT NUMBER

7.1

DETERMINANTS 1
(Second order determinants)

by

A.J.Hobson

UNIT 7.1 - DETERMINANTS 1

SECOND ORDER DETERMINANTS

7.1.1 PAIRS OF SIMULTANEOUS LINEAR EQUATIONS

The subject of Determinants may be introduced by considering, first, a set of two simultaneous linear equations in two unknowns. We shall take them in the form

$$\begin{aligned} a_1x + b_1y + c_1 &= 0, - - - - - - - - - (1) \\ a_2x + b_2y + c_2 &= 0. - - - - - - - - - (2) \end{aligned}$$

If we subtract equation (2) $\times$ b_1 from equation (1) $\times$ b_2, we obtain

$$a_1b_2x - a_2b_1x + c_1b_2 - c_2b_1 = 0.$$

Hence,

$$x = \frac{b_1c_2 - b_2c_1}{a_1b_2 - a_2b_1},$$

provided that $a_1b_2 - a_2b_1 \neq 0$.

Similarly, if we subtract equation (2) $\times$ a_1 from equation (1) $\times$ a_2, we obtain

$$a_2b_1y - a_1b_2y + a_2c_1 - a_1c_2 = 0.$$

Hence,

$$y = -\frac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1},$$

provided that $a_1b_2 - a_2b_1 \neq 0$.

Note:
Other arrangements of the solutions for x and y are possible, but the above arrangements

have been stated for a particular purpose which will be made clear shortly under "Observations".

The Symmetrical Form of the solution

The two separate solutions for x and y may be conveniently written in **"symmetrical"** form as follows:

$$\frac{x}{b_1c_2 - b_2c_1} = \frac{-y}{a_1c_2 - a_2c_1} = \frac{1}{a_1b_2 - a_2b_1},$$

provided that $a_1b_2 - a_2b_1 \neq 0$.

7.1.2 THE DEFINITION OF A SECOND ORDER DETERMINANT

Each of the denominators in the symmetrical form of the previous section has the same general appearance, namely the difference of the products of two pairs of numbers; and we shall rewrite each denominator in a new form using a mathematical symbol called a **"second order determinant"** and defined by the statement:

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = AD - BC.$$

The symbol on the left-hand-side may be called either a second order determinant or a 2×2 determinant; it has two **"rows"** (horizontally), two **"columns"** (vertically) and four **"elements"** (the numbers inside the determinant).

7.1.3 CRAMER'S RULE FOR TWO SIMULTANEOUS LINEAR EQUATIONS

The symmetrical solution to the two simultaneous linear equations may now be written

$$\frac{x}{\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}} = \frac{-y}{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}} = \frac{1}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}},$$

provided that $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0$;

or, in an abbreviated form,

$$\frac{x}{\Delta_1} = \frac{-y}{\Delta_2} = \frac{1}{\Delta_0},$$

provided that $\Delta_0 \neq 0$.

This determinant rule for solving two simultaneous linear equations is called **"Cramer's Rule"** and has equivalent forms for a larger number of equations.

Note:
The interpretation of Cramer's Rule in the case when $a_1b_2 - a_2b_1 = 0$ will be dealt with as a special case after some elementary examples:

Observations

In Cramer's Rule,

1. The determinant underneath x can be remembered by covering up the x terms in the original simultaneous equations, then using the coefficients of y and the constant terms in the pattern which they occupy on the page.
2. The determinant underneath y can be remembered by covering up the y terms in the original simultaneous equations, then using the coefficients of x and the constant terms in the pattern they occupy on the page.
3. The determinant underneath 1 can be remembered by covering up the constant terms in the original simultaneous equations, then using the coefficients of x and y in the pattern they occupy on the page.
4. The final determinant is labelled Δ_0 as a reminder to evaluate it **first**; because, if it happens to be zero, there is no point in evaluating Δ_1 and Δ_2.

EXAMPLES

1. Evaluate the determinant
$$\Delta = \begin{vmatrix} 7 & -2 \\ 4 & 5 \end{vmatrix}.$$

   Solution

$$\Delta = 7 \times 5 - 4 \times (-2) = 35 + 8 = 43$$

2. Express the value of the determinant
$$\Delta = \begin{vmatrix} -p & -q \\ p & -q \end{vmatrix}$$
   in terms of p and q.

Solution

$$\Delta = (-p) \times (-q) - p \times (-q) = p.q + p.q = 2pq.$$

3. Use Cramer's Rule to solve for x and y the simultaneous linear equations

$$\begin{aligned} 5x - 3y &= -3, \\ 2x - y &= -2. \end{aligned}$$

Solution

We may first rearrange the equations in the form

$$\begin{aligned} 5x - 3y + 3 &= 0, \\ 2x - y + 2 &= 0. \end{aligned}$$

Hence, by Cramer's Rule,

$$\frac{x}{\Delta_1} = \frac{-y}{\Delta_2} = \frac{1}{\Delta_0},$$

where

$$\Delta_0 = \begin{vmatrix} 5 & -3 \\ 2 & -1 \end{vmatrix} = -5 + 6 = 1;$$

$$\Delta_1 = \begin{vmatrix} -3 & 3 \\ -1 & 2 \end{vmatrix} = -6 + 3 = -3;$$

$$\Delta_2 = \begin{vmatrix} 5 & 3 \\ 2 & 2 \end{vmatrix} = 10 - 6 = 4.$$

Thus,

$$x = \frac{\Delta_1}{\Delta_0} = -3 \text{ and } y = -\frac{\Delta_2}{\Delta_0} = -4.$$

Special Cases

When using Cramer's Rule, the determinant Δ_0 must not have the value zero. But if it **does** have the value zero, then, the simultaneous linear equations

$$\begin{aligned} a_1x + b_1y + c_1 &= 0, - - - - - - - - - (1) \\ a_2x + b_2y + c_2 &= 0. - - - - - - - - -(2) \end{aligned}$$

are such that

$$a_1b_2 - a_2b_1 = 0.$$

In other words,

$$\frac{a_1}{a_2} = \frac{b_1}{b_2},$$

which means that the x and y terms in one of the equations are proportional to the x and y terms in the other equation.

Two situations may arise which may be illustrated by the following examples:

EXAMPLES

1. For the set of equations

$$\begin{aligned} 3x - 2y &= 5, \\ 6x - 4y &= 10, \end{aligned}$$

   $\Delta_0 = 0$ but the second equation is simply a multiple of the first equation. That is, one of the equations is redundant and so there exists an **infinite number of solutions**; either of the variables may be chosen at random, with the remaining variable being expressible in terms of it.

2. For the set of equations

$$\begin{aligned} 3x - 2y &= 5, \\ 6x - 4y &= 7, \end{aligned}$$

$\Delta_0 = 0$ as before, but there is an inconsistency because, if the second equation is divided by 2, we obtain

$$3x - 2y = 3.5,$$

which is inconsistent with

$$3x - 2y = 5.$$

In this case **there are no solutions at all**.

Summary of the Special Cases

If $\Delta_0 = 0$, further investigation of the simultaneous linear equations is necessary.

7.1.4 EXERCISES

1. Write down the values of the following determinants:

   (a) $\begin{vmatrix} 2 & 3 \\ 1 & 4 \end{vmatrix}$; (b) $\begin{vmatrix} 1 & -2 \\ 2 & 1 \end{vmatrix}$; (c) $\begin{vmatrix} -2 & 3 \\ -1 & 2 \end{vmatrix}$;

   (d) $\begin{vmatrix} x & y \\ y & x \end{vmatrix}$; (e) $\begin{vmatrix} x & x \\ y & y \end{vmatrix}$; (f) $\begin{vmatrix} a & b \\ -b & a \end{vmatrix}$.

2. Use determinants (that is, 'Cramer's Rule') to solve the following sets of simultaneous linear equations:

   (a) $\begin{array}{rcrcr} 19x & + & 6y & = & 39, \\ 13x & - & 8y & = & -6. \end{array}$; (b) $\begin{array}{rcrcrcr} 3x & + & 4y & + & 6 & = & 0, \\ 5x & - & 3y & - & 19 & = & 0. \end{array}$;

   (c) $\begin{array}{rcrcr} 2x & + & 1 & = & 3y, \\ x & - & 5 & = & 7y. \end{array}$; (d) $\begin{array}{rcrcr} 4 & - & 2y & = & x, \\ 7 & + & 3y & = & 2x. \end{array}$;

   (e) $\begin{array}{rcrcr} 3i_1 & + & 2i_2 & = & 5, \\ i_1 & - & 3i_2 & = & 7. \end{array}$; (f) $\begin{array}{rcrcr} 2x & - & 4z & = & 5, \\ x & - & 2z & = & 1. \end{array}$.

3. By expanding out all of the determinants, verify the following results:

   (a) $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = -\begin{vmatrix} b_1 & a_1 \\ b_2 & a_2 \end{vmatrix}$;

   (b) $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 + c_1 & b_1 \\ a_2 + c_2 & b_2 \end{vmatrix} - \begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}$;

   (c) $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 - kb_1 & b_1 \\ a_2 - kb_2 & b_2 \end{vmatrix}$ for any number, k.

7.1.5 ANSWERS TO EXERCISES

1. The values are:
   (a) 5; (b) 5; (c) -1;
   (d) $x^2 - y^2$; (e) 0; (f) $a^2 + b^2$.

2. (a)
$$\frac{x}{-276} = \frac{-y}{621} = \frac{1}{-230};$$
   hence, $x = 1.2$ and $y = 2.7$;
   (b)
$$\frac{x}{-58} = \frac{-y}{-87} = \frac{1}{-29};$$
   hence, $x = 2$ and $y = -3$;
   (c)
$$\frac{x}{22} = \frac{-y}{-11} = \frac{1}{-11};$$
   hence, $x = -2$ and $y = -1$;
   (d)
$$\frac{x}{-26} = \frac{-y}{1} = \frac{1}{-7};$$
   hence, $x = \frac{26}{7}$ and $y = \frac{1}{7}$;
   (e)
$$\frac{i_1}{-29} = \frac{-i_2}{-16} = \frac{1}{-11};$$
   hence, $i_1 = \frac{29}{11}$ and $i_2 = \frac{-16}{11}$;
   (f)
$$\frac{x}{-6} = \frac{-z}{3} = \frac{1}{0},$$
   which means that Cramer's rule breaks down. In fact, the second equation gives two contradictory statements $2x - 4z = 5$ and $2x - 4z = 2$.

3. By following the instructions, the results may be verified.

"JUST THE MATHS"

UNIT NUMBER

7.2

DETERMINANTS 2
(Consistency and third order determinants)

by

A.J.Hobson

UNIT 7.2 - DETERMINANTS 2

CONSISTENCY AND THIRD ORDER DETERMINANTS

7.2.1 CONSISTENCY FOR THREE SIMULTANEOUS LINEAR EQUATIONS IN TWO UNKNOWNS

In a genuine scientific problem involving simultaneous equations, it is not necessarily true that there will be the same number of equations to solve as there are unknowns to be determined. We examine, here, an elementary situation in which there are three equations, but only two unknowns.

Consider the set of equations

$$\begin{aligned} a_1x + b_1y + c_1 &= 0, - - - - - - - - - (1) \\ a_2x + b_2y + c_2 &= 0, - - - - - - - - - (2) \\ a_3x + b_3y + c_3 &= 0, - - - - - - - - - -(3) \end{aligned}$$

where we shall assume that any pair of the three equations has a unique common solution which may be calculated by Cramer's Rule.

In order that the three equations shall be consistent, the common solution of any pair must also satisfy the remaining equation. In particular, the common solution of equations (2) and (3) must also satisfy equation (1).

By Cramer's Rule in equations (2) and (3),

$$\frac{x}{\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}} = \frac{-y}{\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix}} = \frac{1}{\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}}.$$

This solution will satisfy equation (1) provided that

$$a_1\frac{\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}} - b_1\frac{\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}} + c_1 = 0.$$

In other words,

$$a_1\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} = 0.$$

This is the determinant condition for the consistency of three simultaneous linear equations in two unknowns.

7.2.2 THE DEFINITION OF A THIRD ORDER DETERMINANT

In the consistency condition of the previous section, the expression on the left-hand-side is called a **"determinant of the third order"** and is denoted by the symbol

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

It has three **"rows"** (horizontally), three **"columns"** (vertically) and nine **"elements"** (the numbers inside the determinant).

The definition of a third order determinant may be stated in the form

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}.$$

Notes:

(i) Other forms of the definition are also possible and will be encountered in Unit 7.3

(ii) The given formula for evaluating a third order determinant can be remembered by taking each element of the first row in turn and multiplying it by the so-called **"minor"** of the element, which is the second order determinant obtained by covering up the row and column in which the element appears; the results are then combined according to a $+$, $-$, $+$ pattern.

(iii) For the purpose of locating various parts of a determinant, the rows are counted from the top to the bottom and the columns are counted from the left to the right. Each row is read from the left to the right and each column is read from the top to the bottom. Thus, for example, the third element of the second column is b_3.

EXAMPLES

1. Evaluate the determinant

$$\Delta = \begin{vmatrix} -3 & 2 & 7 \\ 0 & 4 & -2 \\ 5 & -1 & 3 \end{vmatrix}.$$

Solution

$$\Delta = -3\begin{vmatrix} 4 & -2 \\ -1 & 3 \end{vmatrix} - 2\begin{vmatrix} 0 & -2 \\ 5 & 3 \end{vmatrix} + 7\begin{vmatrix} 0 & 4 \\ 5 & -1 \end{vmatrix}.$$

That is,

$$\Delta = -3(12-2) - 2(0+10) + 7(0-20) = -190.$$

2. Show that the simultaneous linear equations

$$\begin{aligned} 3x - y + 2 &= 0, \\ 2x + 5y - 1 &= 0, \\ 5x + 4y + 1 &= 0 \end{aligned}$$

are consistent (assuming that any two of the three have a common solution), and obtain the common solution.

Solution

The condition for consistency is that the determinant of coefficients and constants must be zero. We have

$$\begin{vmatrix} 3 & -1 & 2 \\ 2 & 5 & -1 \\ 5 & 4 & 1 \end{vmatrix} = 3\begin{vmatrix} 5 & -1 \\ 4 & 1 \end{vmatrix} - (-1)\begin{vmatrix} 2 & -1 \\ 5 & 1 \end{vmatrix} + 2\begin{vmatrix} 2 & 5 \\ 5 & 4 \end{vmatrix}$$

$$= 3(5+4) + (2+5) + 2(8-25) = 27 + 7 - 34 = 0.$$

Thus, the equations are consistent and, to obtain their common solution, we may solve (say) the first two as follows:

$$\frac{x}{\begin{vmatrix} -1 & 2 \\ 5 & -1 \end{vmatrix}} = \frac{-y}{\begin{vmatrix} 3 & 2 \\ 2 & -1 \end{vmatrix}} = \frac{1}{\begin{vmatrix} 3 & -1 \\ 2 & 5 \end{vmatrix}}.$$

That is,

$$\frac{x}{-9} = \frac{-y}{-7} = \frac{1}{17},$$

which gives

$$x = -\frac{9}{17} \text{ and } y = \frac{7}{17}.$$

Note:
It may be observed that the given set of simultaneous equations above is not an independent set because the third equation happens to be the sum of the other two. We say that the equations are **"linearly dependent"**; and this implies that the rows of the determinant of coefficients and constants are linearly dependent in the same way.
(In this case, Row 3 = Row 1 plus Row 2).

Furthermore, it may shown that the value of a determinant is zero if and only if its rows are linearly dependent. Hence, an alternative way of proving that a set of simultaneous linear equations is a consistent set is to show that they are linearly dependent in some way.

7.2.3 THE RULE OF SARRUS

From the given definition of a third order determinant, the complete **"expansion"** of the determinant may be given, in general, as

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1(b_2c_3 - b_3c_2) - b_1(a_2c_3 - a_3c_2) + c_1(a_2b_3 - a_3b_2)$$

But it may be observed that precisely the same terms may be obtained by first constructing a diagram which consists of the original determinant with the first two columns written out again to the right of this determinant. That is:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \begin{matrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{matrix}.$$

Taking the sum of the possible products of the trios of numbers in the direction $\searrow$ and subtracting the sum of the possible products of the trios of numbers in the $\nearrow$ direction, we obtain the terms

$$(a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3) - (a_3b_2c_1 + b_2c_3a_1 + c_3a_2b_1);$$

and it may be shown that these are exactly the same terms as those obtained by the original formula.

This **"Rule of Sarrus"** makes it possible to evaluate a third order determinant with an electronic calculator, almost without putting pen to paper, provided the calculator memory is used to store, then recall, the various products.

EXAMPLE

$$\begin{vmatrix} -3 & 2 & 7 \\ 0 & 4 & -2 \\ 5 & -1 & 3 \end{vmatrix} = \begin{vmatrix} -3 & 2 & 7 \\ 0 & 4 & -2 \\ 5 & -1 & 3 \end{vmatrix} \begin{matrix} -3 & 2 \\ 0 & 4 \\ 5 & -1 \end{matrix}$$

$$= ([-3].4.3 + 2.[-2].5 + 7.0.[-1]) - (5.4.7 + [-1].[-2].[-3] + 3.0.2)$$

$$= (-36 - 20 + 0) - (140 - 6 + 0) = -56 - 134 = -190$$

as in a previous example.

7.2.4 CRAMER'S RULE FOR THREE SIMULTANEOUS LINEAR EQUATIONS IN THREE UNKNOWNS

In the same way that, under certain conditions, two simultaneous linear equations may be solved by determinants of the second order, it is possible to show that, under certain conditions, three simultaneous linear equations in three unknowns may be solved by determinants of the third order.

The proof of this result will not be included here , but we state it for reference.

The simultaneous linear equations

$$\begin{aligned} a_1x + b_1y + c_1z + d_1 &= 0, \\ a_2x + b_2y + c_2z + d_2 &= 0, \\ a_3x + b_3y + c_3z + d_3 &= 0, \end{aligned}$$

have a common solution, given in symmetrical form, by

$$\frac{x}{\begin{vmatrix} b_1 & c_1 & d_1 \\ b_2 & c_2 & d_2 \\ b_3 & c_3 & d_3 \end{vmatrix}} = \frac{-y}{\begin{vmatrix} a_1 & c_1 & d_1 \\ a_2 & c_2 & d_2 \\ a_3 & c_3 & d_3 \end{vmatrix}} = \frac{z}{\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}} = \frac{-1}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}$$

or

$$\frac{x}{\Delta_1} = \frac{-y}{\Delta_2} = \frac{z}{\Delta_3} = \frac{-1}{\Delta_0},$$

which is called the **"Key"** to the solution and requires that $\Delta_0 \neq 0$.

Again the rule itself is known as **"Cramer's Rule"**.

EXAMPLE

Using the Rule of Sarrus, obtain the common solution of the simultaneous linear equations

$$\begin{aligned} x + 4y - z + 2 &= 0, \\ -x - y + 2z - 9 &= 0, \\ 2x + y - 3z + 15 &= 0. \end{aligned}$$

Solution

The **"Key"** is

$$\frac{x}{\Delta_1} = \frac{-y}{\Delta_2} = \frac{z}{\Delta_3} = \frac{-1}{\Delta_0},$$

where

(i)

$$\Delta_0 = \begin{vmatrix} 1 & 4 & -1 \\ -1 & -1 & 2 \\ 2 & 1 & -3 \end{vmatrix} \begin{matrix} 1 & 4 \\ -1 & -1 \\ 2 & 1 \end{matrix}.$$

Hence,

$$\Delta_0 = (3 + 16 + 1) - (2 + 2 + 12) = 20 - 16 = 4,$$

which is non-zero, and so we may continue:

(ii)

$$\Delta_1 = \left| \begin{array}{ccc|cc} 4 & -1 & 2 & 4 & -1 \\ -1 & 2 & -9 & -1 & 2 \\ 1 & -3 & 15 & 1 & -3 \end{array} \right. .$$

Hence,

$$\Delta_1 = (120 + 9 + 6) - (4 + 108 + 15) = 135 - 127 = 8.$$

(iii)

$$\Delta_2 = \left| \begin{array}{ccc|cc} 1 & -1 & 2 & 1 & -1 \\ -1 & 2 & -9 & -1 & 2 \\ 2 & -3 & 15 & 2 & -3 \end{array} \right. .$$

Hence,

$$\Delta_2 = (30 + 18 + 6) - (8 + 27 + 15) = 54 - 50 = 4.$$

(iv)

$$\Delta_3 = \left| \begin{array}{ccc|cc} 1 & 4 & 2 & 1 & 4 \\ -1 & -1 & -9 & -1 & -1 \\ 2 & 1 & 15 & 2 & 1 \end{array} \right. .$$

Hence,

$$\Delta_3 = (-15 - 72 - 2) - (-4 - 9 - 60) = -89 + 73 = -16.$$

(v) The solutions are therefore

$$x = -\frac{\Delta_1}{\Delta_0} = -\frac{8}{4} = -2;$$

$$y = \frac{\Delta_2}{\Delta_0} = \frac{4}{4} = 1;$$

$$z = -\frac{\Delta_3}{\Delta_0} = -\frac{-16}{4} = 4.$$

Special Cases

If it should happen that $\Delta_0 = 0$ when solving a set of three simultaneous linear equations by Cramer's Rule, earlier work has demonstrated that the rows of Δ_0 must be linearly dependent. That is the three groups of x, y and z terms must be linearly dependent.

Different situations arise according to whether or not the constant terms can also be brought in to the linear dependence relationship and we illustrate with examples as follows:

EXAMPLES

1. For the simultaneous linear equations

$$\begin{aligned} 2x - y + 3z - 5 &= 0, \\ x + 2y - z - 1 &= 0, \\ x - 3y + 4z - 4 &= 0, \end{aligned}$$

the third equation is the difference between the first two and hence it is redundant.

Any solution common to the first two equations will thus be an acceptable solution. In this case, there will be an infinite number of solutions since, for example, we may choose the variable z at random, solving for x and y to obtain

$$x = \frac{11 - 5z}{5} \quad \text{and} \quad y = \frac{5z - 3}{5}.$$

2. For the simultaneous linear equations

$$\begin{aligned} 2x - y + 3z - 5 &= 0, \\ x + 2y - z - 1 &= 0, \\ x - 3y + 4z - 7 &= 0, \end{aligned}$$

the third equation is inconsistent with the difference between the first two equations. That is,

$$x - 3y + 4z - 7 = 0 \text{ is inconsistent with } x - 3y + 4z - 4 = 0.$$

In this case, there are no common solutions.

3. For the simultaneous linear equations

$$\begin{aligned} x - 2y + 3z - 1 &= 0, \\ 2x - 4y + 6z - 2 &= 0, \\ 3x - 6y + 9z - 3 &= 0, \end{aligned}$$

we have only one independent equation since the second and third equations are multiples of the first equation.

Again, there will be an infinite number of solutions which may be obtained by choosing two of the variables at random, then determining the corresponding value of the remaining variable.

Summary of the special cases

If $\Delta_0 = 0$, further investigation of the simultaneous linear equations is necessary.

7.2.5 EXERCISES

1. Show that the simultaneous linear equations

$$\begin{aligned} x + y + 2 &= 0, \\ 3x + 2y - 1 &= 0, \\ 2x + y - 3 &= 0, \end{aligned}$$

are consistent and determine their common solution.

2. Show that the simultaneous linear equations

$$\begin{aligned} 7x - 2y + 1 &= 0, \\ 3x + 2y - 4 &= 0, \\ x - 6y - 9 &= 0, \end{aligned}$$

are inconsistent.

3. Obtain the values of λ for which the simultaneous linear equations

$$\begin{aligned} 3x + 5y + (\lambda - 2) &= 0, \\ 2x + y - 5 &= 0, \\ (\lambda - 1)x + 2y - 10 &= 0, \end{aligned}$$

are consistent.

4. Use Cramer's Rule to solve, for x, y and z, the following simultaneous linear equations:

$$\begin{aligned} 5x + 3y - z + 10 &= 0, \\ -2x - y + 4z - 1 &= 0, \\ -x + 2y - 7z - 17 &= 0. \end{aligned}$$

5. Show that the simultaneous linear equations

$$\begin{aligned} x - y + 7z - 1 &= 0, \\ x + 2y - 3z + 5 &= 0, \\ 5x + 4y + 5z + 13 &= 0 \end{aligned}$$

are linearly dependent and obtain the common solution for which $z = -1$.

7.2.6 ANSWERS TO EXERCISES

1.
$$x = 5 \quad y = -7.$$

2.
$$\Delta_0 \neq 0.$$

3.
$$\lambda = 5 \quad \text{or} \quad \lambda = -23.$$

4.
$$x = -4 \quad y = 3 \quad z = -1.$$

5.
$$x = \frac{8}{3} \quad y = -\frac{16}{3} \quad z = -1.$$

"JUST THE MATHS"

UNIT NUMBER

7.3

DETERMINANTS 3
(Further evaluation of 3 x 3 determinants)

by

A.J.Hobson

UNIT 7.3 - DETERMINANTS 3

FURTHER EVALUATION OF THIRD ORDER DETERMINANTS

7.3.1 EXPANSION BY ANY ROW OR COLUMN

For the numerical evaluation of a third order determinant, the Rule of Sarrus is the easiest rule to apply; but we examine here some alternative versions of the original definition formula which will lead us to important standard properties of determinants.

Let us first re-state the original definition formula as follows:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}.$$

The question naturally arises as to whether the three elements of other rows (or even columns) may multiplied by their minors and the results combined in such a way as to give the same result as in the above formula. It turns out that any row or any column may be used in this way.

In order to illustrate this fact, we state again the more algebraic formula for a third order determinant, namely

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1(b_2c_3 - b_3c_2) - b_1(a_2c_3 - a_3c_2) + c_1(a_2b_3 - a_3b_2).$$

ILLUSTRATION 1 - Expansion by the second row.

It may be observed that the expression

$$-a_2(b_1c_3 - b_3c_1) + b_2(a_1c_3 - a_3c_1) - c_2(a_1b_3 - a_3b_1)$$

gives exactly the same result as in the original formula.

ILLUSTRATION 2 - Expansion by the third column.

It may be observed that the expression

$$c_1(a_2b_3 - a_3b_2) - c_2(a_1b_3 - a_3b_1) + c_3(a_1b_2 - a_2b_1)$$

gives exactly the same result as in the original formula.

Note:
Similar patterns of symbols give the expansions by the remaining rows and columns.

Summary
A third order determinant may be expanded (that is, evaluated) if we multiply each of the three elements in any row or (any column) by its minor then combined the results according the following pattern of so-called **"place-signs"**.

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

Note:
It is useful to have a special name for the **"signed-minor"** which any element of a determinant is mutiplied by, when expanding it by a row or a column. In fact every signed-minor is called a **"cofactor"**. This means that, wherever the place-sign is $+$, the minor and the cofactor are the same; but, wherever the place-sign is $-$, the cofactor is numerically equal to the minor but opposite in sign.

For instance,

(i) The minor of b_1 is $\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix}$,

but the cofactor of b_1 is $-\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix}$.

(ii) The minor and cofactor of b_2 are both equal to $\begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix}$.

7.3.2 ROW AND COLUMN OPERATIONS ON DETERMINANTS

INTRODUCTION

Certain types of problem in scientific work can involve determinants for which some or all of the elements are variable quantities rather than fixed numerical quantities. In these cases, the methods so far encountered for expanding a determinant may not be appropriate.

Described below is a set of standard properties for determinants of any order but, where necessary, they will be explained using either 3×3 determinants or 2×2 determinants.

STANDARD PROPERTIES OF DETERMINANTS

In this section, the methods of expanding a determinant by any row or any column will be useful to have in mind.

1. If all of the elements in a row or a column have the value zero, then the value of the determinant is equal to zero.

   Proof:

   We simply expand the determinant by the row or column of zeros.

2. If all but one of the elements in a row or column are equal to zero, then the value of the determinant is the product of the non-zero element in that row or column with its cofactor.

   Proof:

   We simply expand the determinant by the row or column containing the single non-zero element; and we also notice that the determinant is effectively equivalent to a determinant of one order lower.

   For example,

$$\begin{vmatrix} 5 & 1 & 0 \\ -2 & 4 & 3 \\ 6 & 8 & 0 \end{vmatrix} = -3\begin{vmatrix} 5 & 1 \\ 6 & 8 \end{vmatrix} = -3(40-6) = -102.$$

3. If a determinant contains two identical rows or two identical columns, then the value of the determinant is zero.

   Proof:

   If we expand the determinant by a row or column other than the two identical ones, it will turn out that all of the cofactors have value zero.

   For example,

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{vmatrix} = -4\begin{vmatrix} 2 & 3 \\ 2 & 3 \end{vmatrix} + 5\begin{vmatrix} 1 & 3 \\ 1 & 3 \end{vmatrix} - 6\begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} = 0.$$

4. If two rows, or two columns, are interchanged the value of the determinant is unchanged numerically but it is reversed in sign.

   Proof:

   If we expand the determinant by a row or column other than the two which have been interchanged, then all of the cofactors will be changed in sign.

   For example,

$$\begin{vmatrix} a_1 & c_1 & b_1 \\ a_2 & c_2 & b_2 \\ a_3 & c_3 & b_3 \end{vmatrix} = a_1\begin{vmatrix} c_2 & b_2 \\ c_3 & b_3 \end{vmatrix} - a_2\begin{vmatrix} c_1 & b_1 \\ c_3 & b_3 \end{vmatrix} + a_3\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}.$$

5. If all of the elements in a row or column have a common factor, then this common factor may be removed from the determinant and placed outside.

   Proof:

   Expanding the determinant by the row or column which contains the common factor is equivalent to removing the common factor first, then expanding by the new row or column so created.

   For example,

$$\begin{vmatrix} a_1 & kb_1 & c_1 \\ a_2 & kb_2 & c_2 \\ a_3 & kb_3 & c_3 \end{vmatrix} = k\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix},$$

   if we expand the left-hand-side by the second column.

   Note:

   Another way of stating this property is that, if all of the elements in any row or column of a determinant are multiplied by the same factor, then the value of the determinant is also multiplied by that factor.

6. If the elements of any row in a determinant are altered by adding to them (or subtracting from them) a common multiple of the corresponding elements in another row, then the value of the determinant is unaltered. A similar result applies to columns.

   ILLUSTRATION

   The validity of this result is easily shown in the case of 2×2 determinants as follows:

$$\begin{vmatrix} a_1 + kb_1 & b_1 \\ a_2 + kb_2 & b_2 \end{vmatrix} = [(a_1 + kb_1)b_2 - (a_2 + kb_2)b_1] = a_1b_2 - a_2b_1 = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

The above properties need not normally be used for the evaluation of determinants whose elements are simple numerical values; but, in the examples which follow, we include one such determinant in order to provide a simple introduction to the technique.

We shall use the symbols R_1, R_2 and R_3 to denote Row 1, Row 2 and Row 3; the symbols C_1, C_2 and C_3 will be used to denote Column 1, Column 2 and Column 3; and the symbol $\longrightarrow$ will stand for the word "becomes". The examples use what are called **"row operations"** and **"column operations"**.

EXAMPLES

1. Evaluate the determinant,

$$\begin{vmatrix} 1 & 15 & 7 \\ 2 & 25 & 9 \\ 3 & 10 & 3 \end{vmatrix}$$

Solution

$$\begin{vmatrix} 1 & 15 & 7 \\ 2 & 25 & 9 \\ 3 & 10 & 3 \end{vmatrix} \quad C_1 \longrightarrow C_1 \div 5;$$

$$5\begin{vmatrix} 1 & 3 & 7 \\ 2 & 5 & 9 \\ 3 & 2 & 3 \end{vmatrix} \quad R_2 \longrightarrow R_2 - 2R_1;$$

$$5\begin{vmatrix} 1 & 3 & 7 \\ 0 & -1 & -5 \\ 3 & 2 & 3 \end{vmatrix} \quad R_3 \longrightarrow R_3 - 3R_1;$$

$$\begin{vmatrix} 1 & 3 & 7 \\ 0 & -1 & -5 \\ 0 & -7 & -18 \end{vmatrix}$$

$$= 5(18 - 35) = 5 \times -17 = -85.$$

2. Solve, for x, the equation

$$\begin{vmatrix} x & 5 & 3 \\ 5 & x+1 & 1 \\ -3 & -4 & x-2 \end{vmatrix} = 0.$$

Solution

We could expand the determinant directly, but we would then obtain a cubic equation in x which may not be straightforward to solve.

A better method is to try to obtain factors of this cubic equation **before** expanding the determinant.

It may be observed in this example that the three expressions in each column add up to the same quantity, namely $x + 2$. Thus if we first add Row 2 to Row 1, then add Row 3 to the new Row 1, we shall obtain $x + 2$ as a factor of the first row.

We may write

$$0 = \begin{vmatrix} x & 5 & 3 \\ 5 & x+1 & 1 \\ -3 & -4 & x-2 \end{vmatrix} \quad R_1 \longrightarrow R_1 + R_2 + R_3$$

$$= \begin{vmatrix} x+2 & x+2 & x+2 \\ 5 & x+1 & 1 \\ -3 & -4 & x-2 \end{vmatrix} \quad R_1 \longrightarrow R_1 \div (x+2)$$

$$= (x+2)\begin{vmatrix} 1 & 1 & 1 \\ 5 & x+1 & 1 \\ -3 & -4 & x-2 \end{vmatrix} \quad C_2 \longrightarrow C_2 - C_1 \text{ and } C_3 \longrightarrow C_3 - C_1$$

$$= (x+2)\begin{vmatrix} 1 & 0 & 0 \\ 5 & x-4 & -4 \\ -3 & -1 & x+1 \end{vmatrix}$$

$$= (x+2)[(x-4)(x+1) - 4] = (x-2)(x^2 - 3x - 8).$$

Hence,

$$x = -2 \text{ or } x = \frac{3 \pm \sqrt{9+32}}{2} = \frac{3 \pm \sqrt{41}}{2}.$$

3. Solve, for x, the equation

$$\begin{vmatrix} x-6 & -6 & x-5 \\ 2 & x+2 & 1 \\ 7 & 8 & x+7 \end{vmatrix} = 0,$$

Solution

We may observe that the sum of the corresponding pairs of elements in the first two rows is the same, namely $x - 4$. Hence we may proceed as follows:

$$0 = \begin{vmatrix} x-6 & -6 & x-5 \\ 2 & x+2 & 1 \\ 7 & 8 & x+7 \end{vmatrix} \quad R_1 \longrightarrow R_1 + R_2$$

$$= \begin{vmatrix} x-4 & x-4 & x-4 \\ 2 & x+2 & 1 \\ 7 & 8 & x+7 \end{vmatrix} \quad R_1 \longrightarrow R_1 \div (x-4)$$

$$= (x-4)\begin{vmatrix} 1 & 1 & 1 \\ 2 & x+2 & 1 \\ 7 & 8 & x+7 \end{vmatrix} \quad C_2 \longrightarrow C_2 - C_1 \text{ and } C_3 \longrightarrow C_3 - C_1$$

$$= (x-4)\begin{vmatrix} 1 & 0 & 0 \\ 2 & x & -1 \\ 7 & 1 & x \end{vmatrix}$$

$$= (x-4)(x^2+1).$$

In this case, the only real solution is $x = 4$, the others being complex numbers $x = \pm j$.

4. Solve, for x, the equation

$$\begin{vmatrix} x & 3 & 2 \\ 4 & x+4 & 4 \\ 2 & 1 & x-1 \end{vmatrix},$$

Solution

We may observe that the 2 in Row 1 may be used to reduce to zero the 4 underneath it in Row 2.

Hence,

$$0 = \begin{vmatrix} x & 3 & 2 \\ 4 & x+4 & 4 \\ 2 & 1 & x-1 \end{vmatrix} \quad R_2 \longrightarrow R_2 - 2R_1$$

$$= \begin{vmatrix} x & 3 & 2 \\ 4-2x & x-2 & 0 \\ 2 & 1 & x-1 \end{vmatrix} \quad R_2 \longrightarrow R_2 \div (x-2)$$

$$= (x-2)\begin{vmatrix} x & 3 & 2 \\ -2 & 1 & 0 \\ 2 & 1 & x-1 \end{vmatrix} \quad C_1 \longrightarrow C_1 + 2C_2$$

$$= (x-2)\begin{vmatrix} x+6 & 3 & 2 \\ 0 & 1 & 0 \\ 4 & 1 & x-1 \end{vmatrix}$$

$$= (x-2)[(x+6)(x-1)-8] = (x-2)[x^2+5x-14] = (x-2)(x+7)(x-2).$$

Thus,

$$x = 2 \quad \text{(repeated)} \quad \text{and} \quad x = -7.$$

Note:
It is not possible to cover, by examples, every type of problem which may occur. The secret is first to spend a few seconds examining whether or not the sum or difference of a group of rows or columns can give a common factor immediately. If not, the procedure is to look for ways of obtaining a row or column in which all but one of the elements is zero and hence, effectively, to reduce the order of the determinant.

7.3.3 EXERCISES

1. Use row and/or column operations to evaluate the following determinants:

   (a)
$$\begin{vmatrix} 100 & 101 & 102 \\ 101 & 102 & 103 \\ 102 & 103 & 104 \end{vmatrix};$$

   (b)
$$\begin{vmatrix} 1! & 2! & 3! \\ 2! & 3! & 4! \\ 3! & 4! & 5! \end{vmatrix}.$$

2. Use row and/or column operations to evaluate, in terms of a and b, the determinant
$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1+a & 1 \\ 1 & 1 & 1+b \end{vmatrix}.$$

3. Show that the equation
$$\begin{vmatrix} x & a & b \\ a & x & b \\ a & b & x \end{vmatrix} = 0$$
   has one solution $x = -(a+b)$ and hence solve it completely.

4. Solve completely, for x, the following equations:

   (a)
$$\begin{vmatrix} x-3 & x+2 & x-1 \\ x+2 & x-4 & x \\ x-1 & x+4 & x-5 \end{vmatrix} = 0;$$

   (b)
$$\begin{vmatrix} x+1 & x+2 & 3 \\ 2 & x+3 & x+1 \\ x+3 & 1 & x+2 \end{vmatrix} = 0.$$

7.3.4 ANSWERS TO EXERCISES

1. (a)

$$0$$

(b)

$$24$$

2.

$$ab$$

3.

$$x = -(a+b), \quad x = a, \quad x = b.$$

4. (a)

$$x = \frac{2}{3} \text{ only;}$$

(b)

$$x = -3, \quad x = \pm\sqrt{3}.$$

"JUST THE MATHS"

UNIT NUMBER

7.4

DETERMINANTS 4
(Homogeneous linear equations)

by

A.J.Hobson

UNIT 7.4 - DETERMINANTS 4

HOMOGENEOUS LINEAR EQUATIONS

7.4.1 TRIVIAL AND NON-TRIVIAL SOLUTIONS

This Unit is concerned with a set of simultaneous linear equations in which all of the constant terms have value zero. Most of the discussion will involve three such **"homogeneous"** linear equations of the form

$$\begin{aligned} a_1x + b_1y + c_1z &= 0, \\ a_2x + b_2y + c_2z &= 0, \\ a_3x + b_3y + c_3z &= 0. \end{aligned}$$

These could have been discussed at the same time as Cramer's Rule but are worth considering as a completely separate case since, in scientific applications, they lend themselves conveniently to the methods of row and column operations.

Observations

1. In Cramer's Rule for the above set of equations, if the determinant, Δ_0 of the coefficients of x, y and z is non-zero, there will exist a unique solution, namely $x = 0$, $y = 0$, $z = 0$, since each of the determinants Δ_1, Δ_2 and Δ_3 will contain a column of zeros (that is, the constant terms of the three equations).

   But this solution is obvious from the given set of equations and we call it the **"trivial solution"**.

2. The question arises as to whether it is possible for the set of equations to have any **"non-trivial"** solutions.

3. We shall see that non-trivial solutions occur when the number of equations reduces to less than the number of variables being solved for; that is when the equations are not linearly independent.

   For example, if one of the equations is redundant, we could solve the remaining two in an infinite number of ways by choosing one of the variables at random. Also, if two of the equations are redundant, we could solve the remaining equation in an infinite number of ways by choosing two of the variables at random.

4. It is evident from previous work that the set of homogeneous linear equations will have non-trivial solutions provided that

$$\Delta_0 = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0.$$

5. Once it has been established that non-trivial solutions exist, it can be seen that any solution $x = \alpha$, $y = \beta$, $z = \gamma$ will imply other solutions of the form $x = \lambda\alpha$, $y = \lambda\beta$, $z = \lambda\gamma$, where λ is any non-zero number.

TYPE 1 - One of the three equations is redundant

The non-trivial solutions to a set which reduces to **two** linearly independent homogeneous linear equations in x, y and z may be stated in the form

$$x : y : z = \alpha : \beta : \gamma,$$

in which we mean that

$$\frac{x}{y} = \frac{\alpha}{\beta}, \ \frac{y}{z} = \frac{\beta}{\gamma} \text{ and } \frac{x}{z} = \frac{\alpha}{\gamma}.$$

One method of obtaining these ratios is first to eliminate z between the two equations in order to obtain the ratio $x : y$, then to eliminate y between the two equations in order to find the ratio $x : z$; but a slightly simpler method is described in the first worked example to be discussed shortly.

TYPE 2 - Two of the three equations are redundant

This case arises when the three homogeneous linear equations are multiples of one another.

Again, any solution implies an infinite number of others in the same set of ratios, $x : y : z$. But it turns out that not **all** solutions are in the same set of ratios.

For example, if the only equation remaining is

$$ax + by + cz = 0,$$

we could choose any two of the variables at random and solve for the remaining variable.

In particular, we could substitute $y = 0$ to obtain $x : y : z = -\frac{c}{a} : 0 : 1$;

and we could also substitute $z = 0$ to obtain $x : y : z = -\frac{b}{a} : 1 : 0$

From these two, it is now possible to generate solutions with any value, β, of y and any value, γ, of z (as if we had chosen y and z at random) in order to solve for x.

In fact

$$x = -(\beta.\frac{b}{a} + \gamma.\frac{c}{a}), \quad y = \beta, \quad z = \gamma.$$

Note:
It may be shown that, for a set of homogeneous linear simultaneous equations, no types of solution exist other than those discussed above.

EXAMPLES

1. Show that the homogeneous linear equations

$$\begin{aligned} 2x + y - z &= 0, \\ x - 3y + 2z &= 0, \\ x + 4y - 3z &= 0 \end{aligned}$$

have solutions other than $x = 0$, $y = 0$, $z = 0$ and determine the ratios $x : y : z$ for these non-trivial solutions.

Solution

(a)

$$\Delta_0 = \begin{vmatrix} 2 & 1 & -1 \\ 1 & -3 & 2 \\ 1 & 4 & -3 \end{vmatrix} \begin{matrix} 2 & 1 \\ 1 & -3 \\ 1 & 4 \end{matrix} = (18 + 2 - 4) - (3 + 16 - 3) = 0.$$

Thus, the equations are linearly dependent and, hence, have non-trivial solutions.

Note:
We could, alternatively, have noticed that the first equation is the sum of the second and third equations.

(b) It can always be arranged, in a set of ratios $\alpha : \beta : \gamma$, that any of the quantities which does not have to be equal to zero may be given the value 1. For example, $\frac{\alpha}{\gamma} : \frac{\beta}{\gamma} : 1$ is the same set of ratios as long as γ is not zero.

Let us now suppose that $z = 1$, giving

$$\begin{aligned} 2x + y - 1 &= 0, \\ x - 3y + 2 &= 0, \\ x + 4y - 3 &= 0 \end{aligned}$$

On solving any pair of these equations, we obtain $x = \frac{1}{7}$ and $y = \frac{5}{7}$, which means that

$$x : y : z = \frac{1}{7} : \frac{5}{7} : 1$$

That is,

$$x : y : z = 1 : 5 : 7$$

and any three numbers in these ratios form a solution.

2. Determine the values of λ for which the homogeneous linear equations

$$\begin{aligned} (1-\lambda)x + y - 2z &= 0, \\ -x + (2-\lambda)y + z &= 0, \\ y - (1-\lambda)z &= 0 \end{aligned}$$

have non-trivial solutions.

Solution

First we solve the equation

$$0 = \begin{vmatrix} 1-\lambda & 1 & -2 \\ -1 & 2-\lambda & 1 \\ 0 & 1 & -1-\lambda) \end{vmatrix} \quad R_1 \longrightarrow R_1 - R_3$$

$$= \begin{vmatrix} 1-\lambda & 0 & -1+\lambda \\ -1 & 2-\lambda & 1 \\ 0 & 1 & -1-\lambda \end{vmatrix} \quad R_1 \longrightarrow R_1 \div (1-\lambda)$$

$$= (1-\lambda)\begin{vmatrix} 1 & 0 & -1 \\ -1 & 2-\lambda & 1 \\ 0 & 1 & -1-\lambda \end{vmatrix} \quad C_3 \longrightarrow C_3 + C_1$$

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

$$= -(1-\lambda)(2-\lambda)(1+\lambda)$$

Hence,

$$\lambda = 1, \quad -1 \text{ or } 2.$$

3. Determine the general solution of the homogeneous linear equation

$$3x - 7y + z = 0.$$

Solution

Substituting $y = 0$, we obtain $3x + z = 0$ and hence $x : y : z = -\frac{1}{3} : 0 : 1$.
Substituting $z = 0$, we obtain $3x - 7y = 0$ and hence $x : y : z = \frac{7}{3} : 1 : 0$
The general solution may thus be given by

$$x = \frac{7\beta}{3} - \frac{\gamma}{3}, \quad y = \beta, \quad z = \gamma$$

for arbitrary values of β and γ, though other equivalent versions are possible according to which of the three variables are chosen to have arbitrary values.

7.4.2 EXERCISES

1. Show that the homogeneous linear equations

$$\begin{aligned} x - 2y + 2z &= 0, \\ 2x - 2y - z &= 0, \\ 3x + y + z &= 0 \end{aligned}$$

have no solutions other than the trivial solution.

2. Show that the following sets of homogeneous linear equations have non-trivial solutions and express these solutions as a set of ratios for $x : y : z$

(a)

$$\begin{aligned} x - 2y + z &= 0, \\ x + y - 3z &= 0, \\ 3x - 3y - z &= 0; \end{aligned}$$

(b)

$$\begin{aligned} 3x + y - 2z &= 0, \\ 2x + 4y + 2z &= 0, \\ 4x + 3y - z &= 0. \end{aligned}$$

3. Determine the values of λ for which the homogeneous linear equations

$$\begin{aligned} \lambda x + 2y + 3z &= 0, \\ 2x + (\lambda + 3)y + 6z &= 0, \\ 3x + 4y + (\lambda + 6)z &= 0 \end{aligned}$$

have non-trivial solutions and solve them for the case when λ is an integer.

4. Determine the general solution to the homogeneous linear simultaneous equations

$$\begin{aligned} (\lambda + 1)x - 5y + 3z &= 0, \\ -2x + (\lambda - 8)y + 6z &= 0, \\ -3x - 15y + (\lambda + 11)z &= 0 \end{aligned}$$

in the case when $\lambda = -2$.

7.4.3 ANSWERS TO EXERCISES

1. $$\Delta_0 \neq 0.$$

2. (a)
$$x : y : z = 5 : 4 : 3;$$
(b)
$$x : y : z = 1 : -1 : 1$$

3. $$\lambda = 1,\ 0.83 \ \text{ or } \ -10.83$$
When $\lambda = 1$, $x : y : z = -1 : -1 : 1$

4. $$x = -5\beta + 3\gamma, \quad y = \beta, \quad z = \gamma.$$

"JUST THE MATHS"

UNIT NUMBER

8.1

VECTORS 1
(Introduction to vector algebra)

by

A.J.Hobson

UNIT 8.1 - VECTORS 1 - INTRODUCTION TO VECTOR ALGEBRA

8.1.1 DEFINITIONS

1. A **“scalar”** quantity is one which has magnitude, but is not related to any direction in space.

   Examples: Mass, Speed, Area, Work.

2. A **“vector”** quantity is one which is specified by both a magnitude <u>and</u> a direction in space.

   Examples: Velocity, Weight, Acceleration.

3. A vector quantity with a fixed point of application is called a **“position vector”**.

4. A vector quantity which is restricted to a fixed <u>line</u> of action is called a **“line vector”**.

5. A vector quantity which is defined only by its magnitude and direction is called a **“free vector”**.

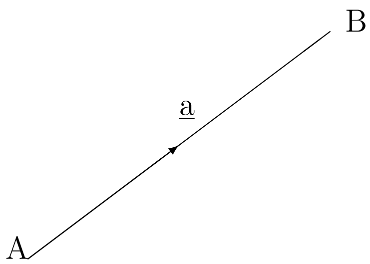
   Note:
   Unless otherwise stated, all vectors in the remainder of these units will be free vectors.

6. A vector quantity can be represented diagramatically by a directed straight line segment in space (with an arrow head) whose direction is that of the vector and whose length represents its magnitude according to a suitable scale.

7. The symbols $\underline{a}$, $\underline{b}$, $\underline{c}$, ... will be used to denote vectors with magnitudes a,b,c.... but it is sometimes more convenient to use a notation such as $\underline{AB}$ which means the vector represented by the line segment drawn from the point A to the point B.

   Notes:
   (i) The magnitude of the vector $\underline{AB}$, which is the length of the line AB can also be denoted by the symbol $|\underline{AB}|$.

   (ii) The magnitude of the vector $\underline{a}$, which is the number a, can also be denoted by the symbol $|\underline{a}|$.



8. A vector whose magnitude is 1 is called a **“unit vector”** and the symbol $\underline{\widehat{a}}$ denotes a unit vector in the same direction as $\underline{a}$. A vector whose magnitude is zero is called a **“zero vector”** and is denoted by **O** or $\underline{O}$. It has indeterminate direction.

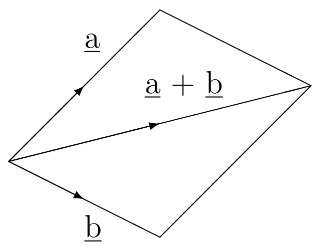
9. Two (free) vectors $\underline{a}$ and $\underline{b}$ are said to be **“equal”** if they have the same magnitude and direction.

   Note:
   This means that two directed straight line segments which are parallel and equal in length may be regarded as representing exactly the same vector.

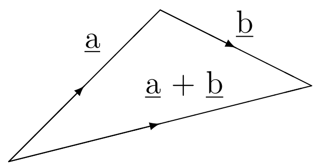
10. A vector whose magnitude is that of $\underline{a}$ but with opposite direction is denoted by $-\underline{a}$.

8.1.2 ADDITION AND SUBTRACTION OF VECTORS

Students may already know how the so-called **"resultant"** (or sum) of particular vectors, like forces, can be determined using either the **"Parallelogram Law"** or alternatively the **"Triangle Law"**. This previous knowledge is not essential here because we now define the sum of two arbitrary vectors diagramatically using either a parallelogram or a triangle. This will then lead also to a definition of subtraction for two vectors.



Parallelogram Law



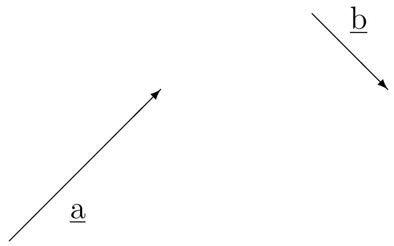
Triangle Law

Notes:
(i) The Triangle Law is more widely used than the Parallelogram Law because of its simplicity. We need to observe that $\underline{a}$ and $\underline{b}$ describe the triangle in the same sense while $\underline{a} + \underline{b}$ describes the triangle in the opposite sense.

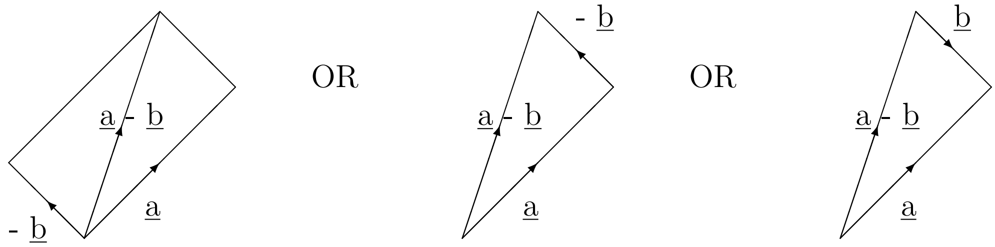
(ii) We define subtraction for vectors by considering that

$$\underline{a} - \underline{b} = \underline{a} + (-\underline{b}).$$

For example, to find $\underline{a} - \underline{b}$ for the vectors $\underline{a}$ and $\underline{b}$ below,



we may construct the following diagrams:



The third figure shows that, to find $\underline{a} - \underline{b}$, we require that $\underline{a}$ and $\underline{b}$ describe the triangle in <u>opposite</u> senses while $\underline{a} - \underline{b}$ describes the triangle in the same sense as $\underline{b}$

(iii) The sum of the three vectors describing the sides of a triangle in the same sense is always the zero vector.

8.1.3 MULTIPLICATION OF A VECTOR BY A SCALAR

If m is any positive real number, $m\underline{a}$ is defined to be a vector in the same direction as $\underline{a}$, but of m times its magnitude.

Similarly $-m\underline{a}$ is a vector in the <u>opposite</u> direction to $\underline{a}$, but of m times its magnitude.

Note:

$\underline{a} = \mathrm{a}\underline{\hat{a}}$ and hence

$$\frac{1}{\mathrm{a}}.\underline{a} = \underline{\hat{a}}.$$

That is, if any vector is multiplied by the reciprocal of its magnitude, we obtain a unit vector in the same direction. This process is called **"normalising the vector"**.

8.1.4 LAWS OF ALGEBRA OBEYED BY VECTORS

(i) The Commutative Law of Addition

$$\underline{a} + \underline{b} = \underline{b} + \underline{a}.$$

(ii) The Associative Law of Addition

$$\underline{a} + (\underline{b} + \underline{c}) = (\underline{a} + \underline{b}) + \underline{c} = \underline{a} + \underline{b} + \underline{c}.$$

(iii) The Associative Law of Multiplication by a Scalar

$$m(n\underline{a}) = (mn)\underline{a} = mn\underline{a}.$$

(iv) The Distributive Laws for Multiplication by a Scalar

$$(m+n)\underline{\mathrm{a}} = m\underline{\mathrm{a}} + n\underline{\mathrm{a}}$$

and

$$m(\underline{\mathrm{a}} + \underline{\mathrm{b}}) = m\underline{\mathrm{a}} + m\underline{\mathrm{b}}.$$

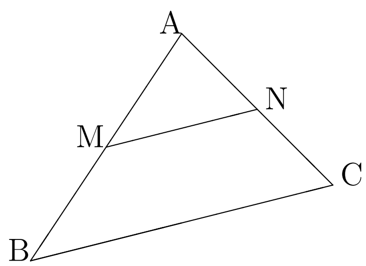
8.1.5 VECTOR PROOFS OF GEOMETRICAL RESULTS

The following examples illustrate how certain geometrical results which could be very cumbersome to prove using traditional geometrical methods can be much more easily proved using a vector method.

EXAMPLES

1. Prove that the line joining the midpoints of two sides of a triangle is parallel to the third side and equal to half of its length.

   Solution



By the Triangle Law,

$$\underline{\mathrm{BC}} = \underline{\mathrm{BA}} + \underline{\mathrm{AC}}$$

and

$$\underline{\mathrm{MN}} = \underline{\mathrm{MA}} + \underline{\mathrm{AN}} = \frac{1}{2}\underline{\mathrm{BA}} + \frac{1}{2}\underline{\mathrm{AC}}.$$

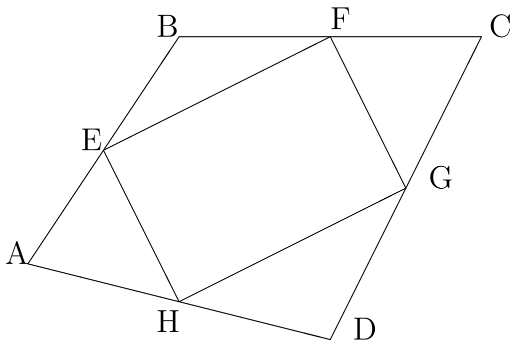
Hence,

$$\underline{\mathrm{MN}} = \frac{1}{2}(\underline{\mathrm{BA}} + \underline{\mathrm{AC}}) = \frac{1}{2}\underline{\mathrm{BC}},$$

which proves the result.

2. ABCD is a quadrilateral (four-sided figure) and E,F,G,H are the midpoints of AB, BC, CD and DA respectively. Show that EFGH is a parallelogram.

Solution



By the Triangle Law,

$$\underline{\mathrm{EF}} = \underline{\mathrm{EB}} + \underline{\mathrm{BF}} = \frac{1}{2}\underline{\mathrm{AB}} + \frac{1}{2}\underline{\mathrm{BC}} = \frac{1}{2}(\underline{\mathrm{AB}} + \underline{\mathrm{BC}}) = \frac{1}{2}\underline{\mathrm{AC}}$$

and also

$$\underline{\mathrm{HG}} = \underline{\mathrm{HD}} + \underline{\mathrm{DG}} = \frac{1}{2}\underline{\mathrm{AD}} + \frac{1}{2}\underline{\mathrm{DC}} = \frac{1}{2}(\underline{\mathrm{AD}} + \underline{\mathrm{DC}}) = \frac{1}{2}\underline{\mathrm{AC}}.$$

Hence,

$$\underline{\mathrm{EF}} = \underline{\mathrm{HG}},$$

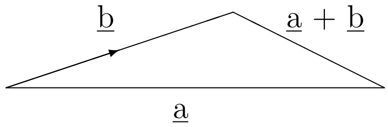
which proves the result.

8.1.6 EXERCISES

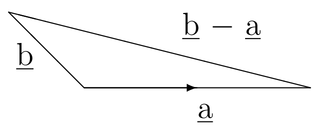
1. Which of the following are vectors and which are scalars ?
   (a) Kinetic Energy; (b) Volume; (c) Force;
   (d) Temperature; (e) Electric Field; (f) Thrust.

2. Fill in the missing arrows for the following vector diagrams:

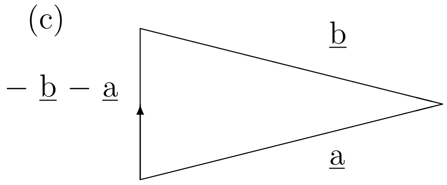
(a)



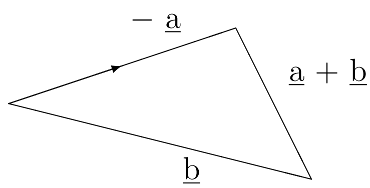
(b)



(c)



(d)



3. ABCDE is a regular pentagon with centre O. Use the Triangle Law of Addition to show that

$$\underline{AB} + \underline{BC} + \underline{CD} + \underline{DE} + \underline{EA} = \mathbf{O}.$$

4. Draw to scale a diagram which illustrates the identity

$$4\underline{a} + 3(\underline{b} - \underline{a}) = \underline{a} + 3\underline{b}.$$

5. $\underline{a}$, $\underline{b}$ and $\underline{c}$ are any three vectors and

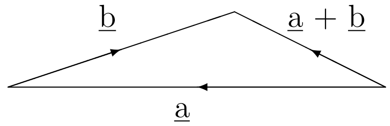
$$\underline{p} = \underline{b} + \underline{c} - 2\underline{a}, \quad \underline{q} = \underline{c} + \underline{a} - 2\underline{b}, \quad \underline{r} = 3\underline{c} - 3\underline{b}.$$

   Show that the vector $3\underline{p} - 2\underline{q}$ is parallel to the vector $5\underline{p} - 6\underline{q} + \underline{r}$.

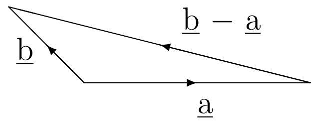
8.1.7 ANSWERS TO EXERCISES

1. (a) Scalar; (b) Scalar; (c) Vector:
   (d) Scalar; (e) Vector; (f) Vector.
2. The completed diagrams are as follows:

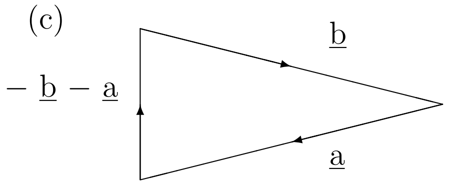
(a)



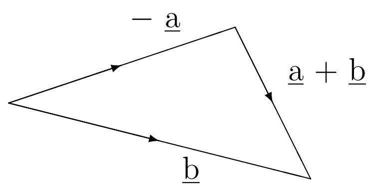
(b)



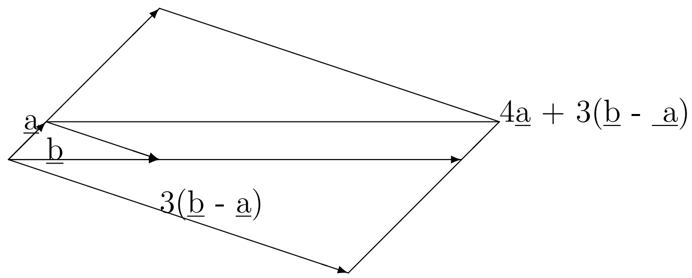
(c)



(d)



3. Join A,B,C,D and E up to the centre, O.
4. The diagram is



5. One vector is a scalar multiple of the other.

"JUST THE MATHS"

UNIT NUMBER

8.2

VECTORS 2
(Vectors in component form)

by

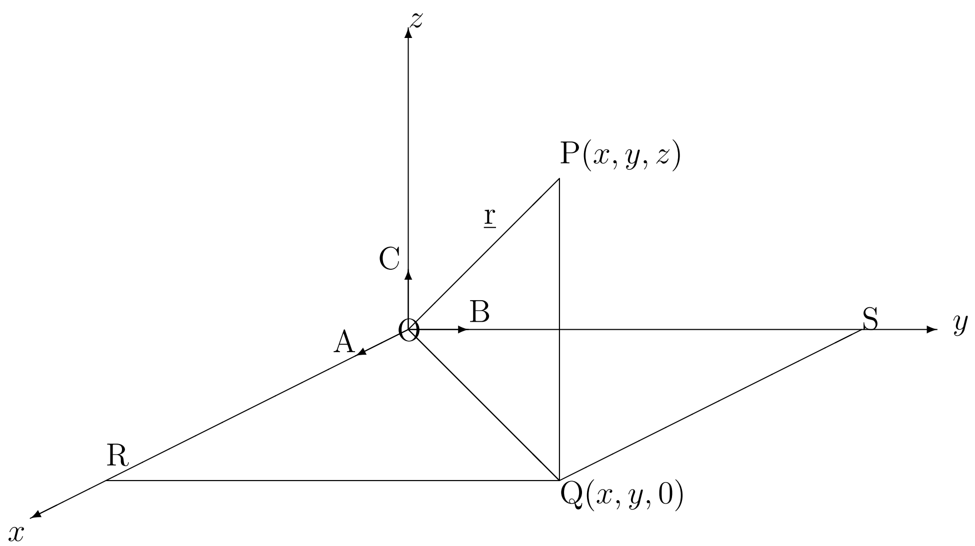
A.J.Hobson

UNIT 8.2 - VECTORS 2 - VECTORS IN COMPONENT FORM

8.2.1 THE COMPONENTS OF A VECTOR

The simplest way to define a vector in space is in terms of **unit vectors** placed along the axes Ox, Oy and Oz of a three-dimensional right-handed cartesian reference system. These unit vectors will be denoted respectively by **i**, **j** and **k** (omitting, for convenience, the "bars" underneath and the "hats" on the top).

Consider the following diagram:



In the diagram, $\underline{\mathrm{OA}} = \mathbf{i}$, $\underline{\mathrm{OB}} = \mathbf{j}$ and $\underline{\mathrm{OC}} = \mathbf{k}$. P is the point with co-ordinates (x, y, z).

By the Triangle Law

$$\underline{\mathrm{r}} = \underline{\mathrm{OP}} = \underline{\mathrm{OQ}} + \underline{\mathrm{QP}} = \underline{\mathrm{OR}} + \underline{\mathrm{RQ}} + \underline{\mathrm{QP}}.$$

That is,

$$\underline{\mathrm{r}} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

Notes:

(i) The fact that we have considered a vector which emanates from the origin is not a special case since we are dealing with free vectors. Nevertheless $\underline{\mathrm{OP}}$ is called the position vector of the point P.

(ii) The numbers x, y and z are called the **"components"** of $\underline{\mathrm{OP}}$ (or of any other vector in space with the same magnitude and direction as $\underline{\mathrm{OP}}$).

(iii) To multiply (or divide) a vector in component form by a scalar, we simply multiply (or divide) each of its components by that scalar.

8.2.2 THE MAGNITUDE OF A VECTOR IN COMPONENT FORM

Referring to the diagram in section 8.2.1, Pythagoras' Theorem gives

$$(\mathrm{OP})^2 = (\mathrm{OQ})^2 + (\mathrm{QP})^2 = (\mathrm{OR})^2 + (\mathrm{RQ})^2 + (\mathrm{QP})^2.$$

That is,

$$\mathrm{r} = |x\mathbf{i} + y\mathbf{j} + z\mathbf{k}| = \sqrt{x^2 + y^2 + z^2}.$$

EXAMPLE

Determine the magnitude of the vector

$$\underline{\mathrm{a}} = 5\mathbf{i} - 2\mathbf{j} + \mathbf{k}$$

and hence obtain a unit vector in the same direction.

Solution

$$|\underline{\mathrm{a}}| = \mathrm{a} = \sqrt{5^2 + (-2)^2 + 1^2} = \sqrt{30}.$$

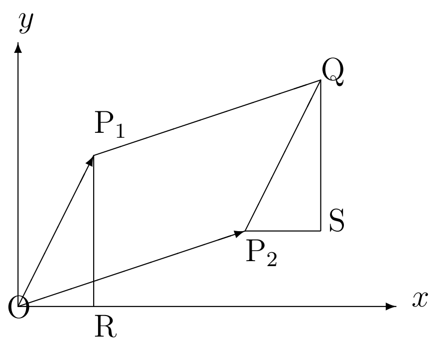
Hence, a unit vector in the same direction as $\underline{\mathrm{a}}$ is obtained by normalising $\underline{\mathrm{a}}$; that is, dividing it by its own magnitude.

The required unit vector is

$$\widehat{\underline{\mathrm{a}}} = \frac{1}{\mathrm{a}}.\underline{\mathrm{a}} = \frac{5\mathbf{i} - 2\mathbf{j} + \mathbf{k}}{\sqrt{30}}.$$

8.2.3 THE SUM AND DIFFERENCE OF VECTORS IN COMPONENT FORM

We consider, first, a situation in **two** dimensions where two vectors are added together.



In the diagram, suppose P_1 has co-ordinates (x_1, y_1) and suppose P_2 has co-ordinates (x_2, y_2).

Then, since the triangle ORP_1 has exactly the same shape as the triangle P_2SQ, the co-ordinates of Q must be $(x_1 + x_2, y_1 + y_2)$.

But, by the Parallelogram Law, $\underline{OQ}$ is the sum of $\underline{OP_1}$ and $\underline{OP_2}$.

That is,

$$(x_1\mathbf{i} + y_1\mathbf{j}) + (x_2\mathbf{i} + y_2\mathbf{j}) = (x_1 + x_2)\mathbf{i} + (y_1 + y_2)\mathbf{j},$$

showing that the sum of two vectors may be found by adding together their separate components.

It can be shown that this result applies in three dimensions also and that, to find the **difference** of two vectors, we calculate the difference of their separate components.

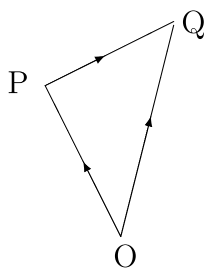
EXAMPLE

Two points P and Q in space have cartesian co-ordinates $(-3, 1, 4)$ and $(2, -2, 5)$ respectively. Determine the vector $\underline{PQ}$.

Solution

We are given that

$$\underline{OP} = -3\mathbf{i} + \mathbf{j} + 4\mathbf{k} \;\text{ and }\; \underline{OQ} = 2\mathbf{i} - 2\mathbf{j} + 5\mathbf{k}.$$



By the triangle Law,

$$\underline{PQ} = \underline{OQ} - \underline{OP} = 5\mathbf{i} - 3\mathbf{j} + \mathbf{k}.$$

Note:
The vector $\underline{PQ}$ is, of course, the vector drawn from the point P to the point Q and it may seem puzzling that the result just obtained appears to be a vector drawn from the origin to the point $(5, -3, 1)$. However, we need to use again the fact that we are dealing with free vectors and the vector drawn from the origin to the point $(5, -3, 1)$ is parallel and equal in length to $\underline{PQ}$; in other words, it is the **same** as $\underline{PQ}$.

8.2.4 THE DIRECTION COSINES OF A VECTOR

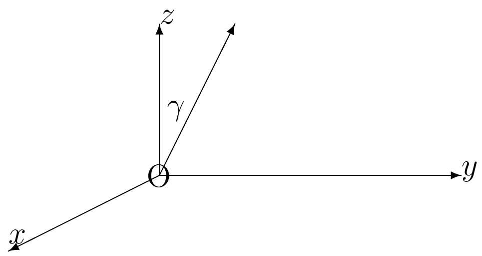
Suppose that

$$\underline{\mathrm{OP}} = \underline{\mathrm{r}} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$$

and suppose that $\underline{\mathrm{OP}}$ makes angles α, β and γ with Ox, Oy and Oz respectively.

Then,

$$\cos\alpha = \frac{x}{\mathrm{r}}, \quad \cos\beta = \frac{y}{\mathrm{r}} \text{ and } \cos\gamma = \frac{z}{\mathrm{r}}.$$



The three quantities $\cos\alpha$, $\cos\beta$ and $\cos\gamma$ are called the **"direction cosines"** of $\underline{\mathrm{r}}$.

Any three numbers in the same ratio as the direction cosines are said to form a set of **"direction ratios"** for the vector $\underline{\mathrm{r}}$ and we note that $x : y : z$ is one possible set of direction ratios.

EXAMPLE

The direction cosines of the vector

$$6\mathbf{i} + 2\mathbf{j} - \mathbf{k}$$

are

$$\frac{6}{\sqrt{41}}, \ \frac{2}{\sqrt{41}} \text{ and } \frac{-1}{\sqrt{41}},$$

since the vector has magnitude $\sqrt{36+4+1} = \sqrt{41}$.

A set of direction ratios for this vector are $6 : 2 : -1$.

8.2.5 EXERCISES

1. The position vectors of two points P and Q are, respectively,

$$\underline{\mathrm{r}_1} = 2\mathbf{i} + 3\mathbf{j} - \mathbf{k} \text{ and } \underline{\mathrm{r}_2} = 4\mathbf{i} - 3\mathbf{j} + 2\mathbf{k}.$$

Determine the vector $\underline{\mathrm{PQ}}$ in terms of **i** , **j** and **k** and hence obtain

(a) its magnitude;

(b) its direction cosines.

2. Obtain a unit vector which is parallel to the vector $\underline{\mathrm{a}} = 3\mathbf{i} - \mathbf{j} + 5\mathbf{k}$.

3. If $\underline{\mathrm{a}} = 3\mathbf{i} - \mathbf{j} - 4\mathbf{k}$, $\underline{\mathrm{b}} = -2\mathbf{i} + 4\mathbf{j} - 3\mathbf{k}$ and $\underline{\mathrm{c}} = \mathbf{i} + 2\mathbf{j} - \mathbf{k}$, determine the following:

(a)

$$2\underline{\mathrm{a}} - \underline{\mathrm{b}} + 3\underline{\mathrm{c}};$$

(b)

$$|\underline{\mathrm{a}} + \underline{\mathrm{b}} + \underline{\mathrm{c}}|;$$

(c)

$$|3\underline{\mathrm{a}} - 2\underline{\mathrm{b}} + 4\underline{\mathrm{c}}|;$$

(d) a unit vector which is parallel to

$$3\underline{\mathrm{a}} - 2\underline{\mathrm{b}} + 4\underline{\mathrm{c}}.$$

4. Prove that the vectors $\underline{\mathrm{a}} = 3\mathbf{i} + \mathbf{j} - 2\mathbf{k}$, $\underline{\mathrm{b}} = -\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}$ and $\underline{\mathrm{c}} = 4\mathbf{i} - 2\mathbf{j} - 6\mathbf{k}$ can form the sides of a triangle.

Determine also the lengths of the **"medians"** of this triangle (that is, the lines joining each vertex to the mid-point of the opposite side).

8.2.6 ANSWERS TO EXERCISES

1. $\underline{\mathrm{PQ}} = 2\mathbf{i} - 6\mathbf{j} + 3\mathbf{k}$.

   (a) $|\underline{\mathrm{PQ}}| = 7$;

   (b) The direction cosines are $\frac{2}{7}$, $-\frac{6}{7}$ and $\frac{3}{7}$.

2.

$$\pm\frac{3\mathbf{i} - \mathbf{j} + 5\mathbf{k}}{\sqrt{35}}.$$

3. (a)

$$11\mathbf{i} - 8\mathbf{k};$$

   (b)

$$\sqrt{93};$$

   (c)

$$\sqrt{398};$$

   (d)

$$\pm\frac{17\mathbf{i} - 3\mathbf{j} - 10\mathbf{k}}{\sqrt{398}}.$$

4. $\underline{\mathrm{a}} = \underline{\mathrm{b}} + \underline{\mathrm{c}}$; therefore the vectors form a triangle.

   The medians have lengths equal to

$$5\sqrt{\frac{3}{2}},\ \sqrt{6} \text{ and } \frac{\sqrt{114}}{2}.$$

"JUST THE MATHS"

UNIT NUMBER

8.3

VECTORS 3
(Multiplication of one vector by another)

by

A.J.Hobson

UNIT 8.3 - VECTORS 3

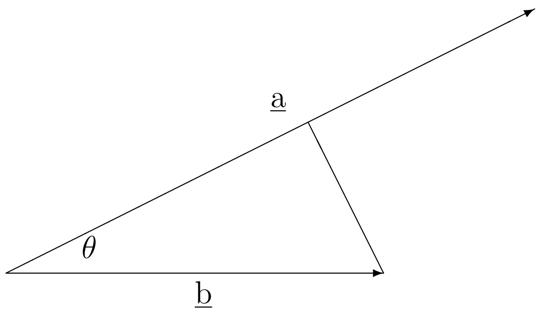
MULTIPLICATION OF ONE VECTOR BY ANOTHER

8.3.1 THE SCALAR PRODUCT (or "Dot" Product)

DEFINITION

The **"Scalar Product"** of two vectors $\underline{a}$ and $\underline{b}$ is defined as $ab\cos\theta$, where θ is the angle between the directions of $\underline{a}$ and $\underline{b}$, drawn so that they have a common end-point and are directed away from that point. The Scalar Product is denoted by $\underline{a} \bullet \underline{b}$ so that

$$\underline{a} \bullet \underline{b} = ab\cos\theta$$



$$0 \le \theta \le \pi$$

Scientific Application

If $\underline{b}$ were a force of magnitude b, then $b\cos\theta$ would be its resolution (or component) along the vector $\underline{a}$. Hence, $\underline{a} \bullet \underline{b}$ would represent the work done by $\underline{b}$ in moving an object along the vector $\underline{a}$. Similarly, if $\underline{a}$ were a force of magnitude a, then $a\cos\theta$ would be its resolution (or component) along the vector $\underline{b}$. Hence, $\underline{a} \bullet \underline{b}$ would represent the work done by $\underline{a}$ in moving an object along the vector $\underline{b}$.

8.3.2 DEDUCTIONS FROM THE DEFINITION OF DOT PRODUCT

(i) $\underline{a} \bullet \underline{a} = a^2$.

Proof:

Clearly, the angle between $\underline{a}$ and itself is zero so that

$$\underline{a} \bullet \underline{a} = a.a\cos 0 = a^2.$$

(ii) $\underline{a} \bullet \underline{b}$ can be interpreted as the magnitude of one vector times the perpendicular projection of the other vector onto it.

Proof:

$b\cos\theta$ is the perpendicular projection of $\underline{b}$ onto $\underline{a}$ and $a\cos\theta$ is the perpendicular projection of $\underline{a}$ onto $\underline{b}$.

(iii) $\underline{a} \bullet \underline{b} = \underline{b} \bullet \underline{a}$.

Proof:

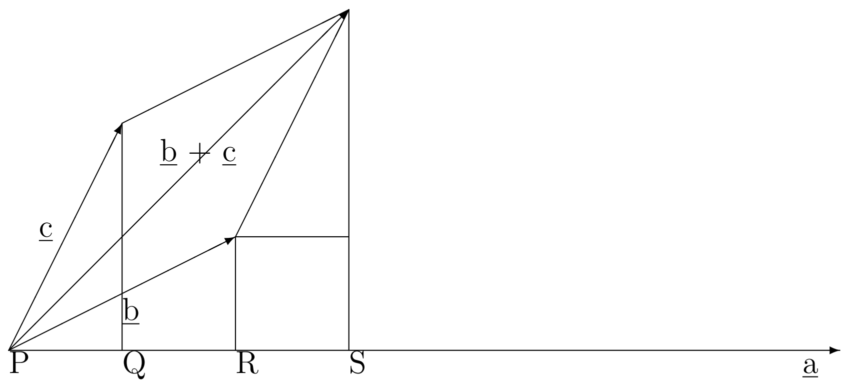
This follows since $\text{ab}\cos\theta = \text{ba}\cos\theta$.

(iv) Two non-zero vectors are perpendicular if and only if their Scalar Product is zero.

Proof:

$\underline{a}$ is perpendicular to $\underline{b}$ if and only if the angle $\theta = \frac{\pi}{2}$; that is, if and only if $\cos\theta = 0$ and hence, $\text{ab}\cos\theta = 0$.

(v) $\underline{a} \bullet (\underline{b} + \underline{c}) = \underline{a} \bullet \underline{b} + \underline{a} \bullet \underline{c}$.



The result follows from (ii) since the projections PR and PQ of $\underline{b}$ and $\underline{c}$ respectively onto $\underline{a}$ add up to the projection PS of $\underline{b} + \underline{c}$ onto $\underline{a}$.

Note:
We need to observe that RS is equal in length to PQ.

(vi) The Scalar Product of any two of the standard unit vectors **i**, **j** and **k** is given by the following multiplication table:

$\bullet$	**i**	**j**	**k**
i	1	0	0
j	0	1	0
k	0	0	1

That is, $\mathbf{i} \bullet \mathbf{i} = 1$, $\mathbf{j} \bullet \mathbf{j} = 1$ and $\mathbf{k} \bullet \mathbf{k} = 1$;

but,

$\mathbf{i} \bullet \mathbf{j} = 0$, $\mathbf{i} \bullet \mathbf{k} = 0$ and $\mathbf{j} \bullet \mathbf{k} = 0$.

8.3.3 THE STANDARD FORMULA FOR DOT PRODUCT

If

$$\underline{\mathrm{a}} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} \quad \text{and} \quad \underline{\mathrm{b}} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k},$$

then

$$\underline{\mathrm{a}} \;\bullet\; \underline{\mathrm{b}} = a_1b_1 + a_2b_2 + a_3b_3.$$

Proof:

This result follows easily from the multiplication table in (vi).

Note: The angle between two vectors

If θ is the angle between the two vectors $\underline{\mathrm{a}}$ and $\underline{\mathrm{b}}$, then

$$\cos\theta = \frac{\underline{\mathrm{a}} \bullet \underline{\mathrm{b}}}{\mathrm{ab}}.$$

Proof:

This result is just a restatement of the original definition of a Scalar Product.

EXAMPLE

If

$$\underline{\mathrm{a}} = 2\mathbf{i} + 2\mathbf{j} - \mathbf{k} \quad \text{and} \quad \underline{\mathrm{b}} = 3\mathbf{j} - 4\mathbf{k},$$

then,

$$\underline{\mathrm{a}} \bullet \underline{\mathrm{b}} = 2 \times 0 + 2 \times 3 + (-1) \times (-4) = 10.$$

Hence,

$$\cos\theta = \frac{10}{\sqrt{2^2 + 2^2 + 1^2}\sqrt{3^2 + 4^2}} = \frac{10}{15} = \frac{2}{3}.$$

Thus,

$$\theta = 48.19^\circ \quad \text{or} \quad 0.84 \text{ radians}.$$

8.3.4 THE VECTOR PRODUCT (or "Cross" Product)

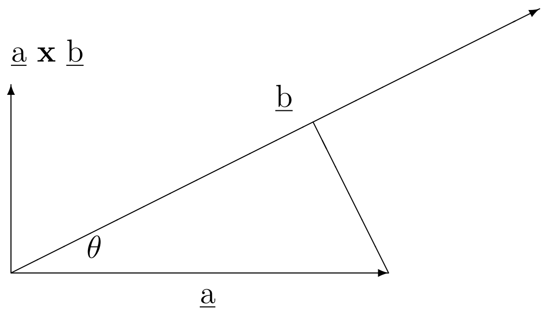
DEFINITION

If θ is the angle between two vectors $\underline{a}$ and $\underline{b}$, drawn so that they have a common end-point and are directed away from that point, then the **"Vector Product"** of $\underline{a}$ and $\underline{b}$ is defined to be a vector of magnitude

$$ab \sin\theta,$$

in a direction which is perpendicular to the plane containing $\underline{a}$ and $\underline{b}$ and in a sense which obeys the **"right-hand-thread screw rule"** in turning from $\underline{a}$ to $\underline{b}$. The Vector Product is denoted by

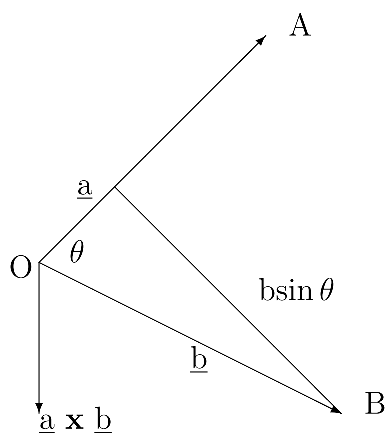
$$\underline{a} \mathbf{x} \underline{b}.$$



$$0 \leq \theta \leq \pi$$

Scientific Application

Consider the following diagram:



Suppose that the vector $\underline{OA} = \underline{a}$ represents a force acting at the point O and that the vector $\underline{OB} = \underline{b}$ is the position vector of the point B. Let the angle between the two vectors be θ.

Then the **"moment"** of the force $\underline{OA}$ about the point B is a vector whose magnitude is

$$ab \sin\theta$$

and whose direction is perpendicular to the plane of O, A and B in a sense which obeys the right-hand-thread screw rule in turning from $\underline{OA}$ to $\underline{OB}$. That is

$$\text{Moment} = \underline{a} \ \mathbf{x} \ \underline{b}.$$

Note:
The quantity $b \sin\theta$ is the perpendicular distance from the point B to the force $\underline{\mathrm{OA}}$.

8.3.5 DEDUCTIONS FROM THE DEFINITION OF CROSS PRODUCT

(i)
$$\underline{a} \textbf{ x } \underline{b} = -(\underline{b} \textbf{ x } \underline{a}) = (-\underline{b}) \textbf{ x } \underline{a} = \underline{b} \textbf{ x } (-\underline{a}).$$

Proof:
This follows easily by considering the implications of the right-hand-thread screw rule.

(ii) Two vectors are parallel if and only if their Cross Product is a zero vector.

Proof:
Two vectors are parallel if and only if the angle, θ, between them is zero or π. In either case, $\sin\theta = 0$, which means that $ab\sin\theta = 0$; that is, $|\underline{a} \textbf{ x } \underline{b}| = 0$.

(iii) The Cross Product of a vector with itself is a zero vector.

Proof:
Clearly, the angle between a vector, $\underline{a}$, and itself is zero. Hence,
$$|\underline{a} \textbf{ x } \underline{a}| = \mathrm{a.a.}\sin 0 = 0.$$

(iv)
$$\underline{a} \textbf{ x } (\underline{b} + \underline{c}) = \underline{a} \textbf{ x } \underline{b} + \underline{a} \textbf{ x } \underline{c}.$$

Proof:
This is best proved using the standard formula for a Cross Product in terms of components (see 8.3.6 below).

(v) The multiplication table for the Cross Products of the standard unit vectors **i**, **j** and **k** is as follows:

x	i	j	k
i	O	k	–j
j	– k	O	i
k	j	– i	O

That is,
i x i = **O**, **j x j** = **O**, **k x k** = **O**, **i x j** = **k**, **j x k** = **i**, **k x i** = **j**, **j x i** = – **k**, **k x j** = – **i** and **i x k** = – **j**.

8.3.6 THE STANDARD FORMULA FOR CROSS PRODUCT

If
$$\underline{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} \quad \text{and} \quad \underline{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k},$$
then,
$$\underline{a} \textbf{ x } \underline{b} = (a_2b_3 - a_3b_2)\mathbf{i} - (a_1b_3 - a_3b_1)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k}.$$

This is usually abbreviated to

$$\underline{a}\ \mathbf{x}\ \underline{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix},$$

the symbol on the right hand side being called a **"determinant"** (see Unit 7.2).

EXAMPLES

1. If $\underline{a} = 2\mathbf{i} + 2\mathbf{j} - \mathbf{k}$ and $\underline{b} = 3\mathbf{j} - 4\mathbf{k}$, determine $\underline{a}\ \mathbf{x}\ \underline{b}$.

   Solution

$$\underline{a}\ \mathbf{x}\ \underline{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 2 & -1 \\ 0 & 3 & -4 \end{vmatrix} = (-8+3)\mathbf{i} - (-8-0)\mathbf{j} + (6-0)\mathbf{k} = -5\mathbf{i} + 8\mathbf{j} + 6\mathbf{k}.$$

2. Show that, for any two vectors $\underline{a}$ and $\underline{b}$,

$$(\underline{a} + \underline{b})\ \mathbf{x}\ (\underline{a} - \underline{b}) = 2(\underline{b}\ \mathbf{x}\ \underline{a}).$$

   Solution

   The left hand side =

$$\underline{a}\ \mathbf{x}\ \underline{a} - \underline{a}\ \mathbf{x}\ \underline{b} + \underline{b}\ \mathbf{x}\ \underline{a} - \underline{b}\ \mathbf{x}\ \underline{b}.$$

   That is,

$$\mathbf{O} + \underline{b}\ \mathbf{x}\ \underline{a} + \underline{b}\ \mathbf{x}\ \underline{a} = 2(\underline{b}\ \mathbf{x}\ \underline{a}).$$

3. Determine the area of the triangle defined by the vectors

$$\underline{a} = \mathbf{i} + \mathbf{j} + \mathbf{k} \quad \text{and} \quad \underline{b} = 2\mathbf{i} - 3\mathbf{j} + \mathbf{k}.$$

   Solution

   If θ is the angle between the two vectors $\underline{a}$ and $\underline{b}$, then the area of the triangle is $\frac{1}{2}\mathrm{ab}\sin\theta$ from elementary trigonometry. The area is therefore given by

$$\frac{1}{2}|\underline{a}\ \mathbf{x}\ \underline{b}|.$$

   That is,

$$\text{Area} = \frac{1}{2}\left\| \begin{matrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 1 & 1 \\ 2 & -3 & 1 \end{matrix} \right\| = \frac{1}{2}|4\mathbf{i} + \mathbf{j} - 5\mathbf{k}|.$$

   This gives

$$\text{Area} = \frac{1}{2}\sqrt{16+1+25} = \frac{1}{2}\sqrt{42} \simeq 3.24$$

8.3.7 EXERCISES

1. In the following cases, evaluate the Scalar Product $\underline{a} \bullet \underline{b}$ and hence determine the angle, θ between $\underline{a}$ and $\underline{b}$:

   (a) $\underline{a} = 2\mathbf{i} + 2\mathbf{j} - \mathbf{k}$ and $\underline{b} = 3\mathbf{i} - 6\mathbf{j} + 2\mathbf{k}$;

   (b) $\underline{a} = \mathbf{i} + 2\mathbf{j} - \mathbf{k}$ and $\underline{b} = 3\mathbf{j} + \mathbf{k}$;

   (c) $\underline{a} = -\mathbf{i} - \mathbf{j} + 4\mathbf{k}$ and $\underline{b} = 7\mathbf{i} - 2\mathbf{k}$.

2. Find out which of the following pairs of vectors are perpendicular and determine the cosine of the angle between those which are not:

   (a) $3\mathbf{j}$ and $2\mathbf{j} - 2\mathbf{k}$;

   (b) $\mathbf{i} + 3\mathbf{j} - 5\mathbf{k}$ and $-\mathbf{i} + 2\mathbf{j} + \mathbf{k}$;

   (c) $2\mathbf{i} + 10\mathbf{k}$ and $7\mathbf{j}$;

   (d) $2\mathbf{i} + 2\mathbf{j} - \mathbf{k}$ and $6\mathbf{i} - 3\mathbf{j} + 2\mathbf{k}$.

3. If $\underline{a} = 2\mathbf{i} + \mathbf{j} + \mathbf{k}$, $\underline{b} = \mathbf{i} - 2\mathbf{j} - 2\mathbf{k}$ and $\underline{c} = 3\mathbf{i} - 4\mathbf{j} + 2\mathbf{k}$, determine the length of the projection of $\underline{a} + \underline{c}$ onto $\underline{b}$.

4. If $\underline{a} = 2\mathbf{i} - 3\mathbf{j} + 5\mathbf{k}$ and $\underline{b} = 3\mathbf{i} + \mathbf{j} - 2\mathbf{k}$, evaluate

$$(\underline{a} + \underline{b}) \bullet (\underline{a} - \underline{b}).$$

5. Determine the components of the vector $\underline{a}$ **x** $\underline{b}$ in the following cases:

   (a) $\underline{a} = 2\mathbf{i} + 2\mathbf{j} - \mathbf{k}$ and $\underline{b} = 3\mathbf{i} - 6\mathbf{j} + 2\mathbf{k}$;

   (b) $\underline{a} = \mathbf{i} + 2\mathbf{j} - \mathbf{k}$ and $\underline{b} = 3\mathbf{j} + \mathbf{k}$;

   (c) $\underline{a} = -\mathbf{i} - \mathbf{j} + 4\mathbf{k}$ and $\underline{b} = 7\mathbf{i} - 2\mathbf{k}$.

6. If $\underline{a} = 3\mathbf{i} - \mathbf{j} + 2\mathbf{k}$ and $\underline{b} = \mathbf{i} + 3\mathbf{j} - 2\mathbf{k}$, show that $\underline{a}$ **x** $\underline{b}$ is perpendicular to the vector $\underline{c} = 9\mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$.

7. Given that $\underline{a}$ **x** $\underline{b}$ is perpendicular to each one of the vectors $\underline{a}$ and $\underline{b}$, determine a unit vector which is perpendicluar to each one of the vectors $\underline{a} = 2\mathbf{i} - \mathbf{j} + \mathbf{k}$ and $\underline{b} = 3\mathbf{i} + 4\mathbf{j} - \mathbf{k}$.
   Calculate also the sine of the angle, θ, between $\underline{a}$ and $\underline{b}$.

8. Determine the area of the triangle whose vertices are the points A(3, − 1,2), B(1, − 1, − 3) and C(4, − 3,1) in space. State your answer correct to two places of decimals.

8.3.8 ANSWERS TO EXERCISES

1. (a) Scalar Product $= -8$, $\cos\theta \simeq -0.381$ and $\theta \simeq 112.4°$;
   (b) Scalar Product $= 5$, $\cos\theta \simeq 0.645$ and $\theta \simeq 49.80°$;
   (c) Scalar Product $= -15$, $\cos\theta \simeq -0.485$ and $\theta \simeq 119.05°$
2. (a) Cosine $\simeq 0.707$;
   (b) The vectors are perpendicular;
   (c) The vectors are perpendicular;
   (d) Cosine $\simeq 0.190$
3. The length of the projection is $\frac{5}{3}$.
4. The value of the Dot Product is 24.
5. (a) The components are $-2, -7, -18$;
   (b) The components are $5, -1, 3$;
   (c) The components are $2, 26, 7$.
6. Show that $(\underline{a}\ \mathbf{x}\ \underline{b}) \bullet \underline{c} = 0$.
7. A unit vector is

$$\pm\frac{-3\mathbf{i} + 5\mathbf{j} + 11\mathbf{k}}{\sqrt{155}}$$

   and

$$\sin\theta = \frac{\sqrt{155}}{\sqrt{6}.\sqrt{26}} \simeq 0.997$$

8. The area is 6.42

"JUST THE MATHS"

UNIT NUMBER

8.4

VECTORS 4
(Triple products)

by

A.J.Hobson

UNIT 8.4 - VECTORS 4

TRIPLE PRODUCTS

INTRODUCTION

Once the ideas of scalar (dot) product and vector (cross) product for two vectors has been introduced, it is then possible to consider certain products of three or more vectors where, in some cases, there may be a mixture of scalar and vector products.

8.4.1 THE TRIPLE SCALAR PRODUCT

DEFINITION 1

Given three vectors $\underline{a}$, $\underline{b}$ and $\underline{c}$, expressions such as

$$\underline{a} \bullet (\underline{b} \mathbf{x} \underline{c}), \quad \underline{b} \bullet (\underline{c} \mathbf{x} \underline{a}), \quad \underline{c} \bullet (\underline{a} \mathbf{x} \underline{b})$$

or

$$(\underline{a} \mathbf{x} \underline{b}) \bullet \underline{c}, \quad (\underline{b} \mathbf{x} \underline{c}) \bullet \underline{a}, \quad (\underline{c} \mathbf{x} \underline{a}) \bullet \underline{b}$$

are called **"triple scalar products"** because their results are all scalar quantities. Strictly speaking, the brackets are not necessary because there is no ambiguity without them; that is, it is not possible to form the vector product of a vector with the result of a scalar product.

In the work which follows, we shall take $\underline{a} \bullet (\underline{b} \mathbf{x} \underline{c})$ as the typical triple scalar product.

The formula for a triple scalar product

Suppose that

$$\underline{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}, \quad \underline{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k} \text{ and } \underline{c} = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}.$$

Then,

$$\underline{a} \bullet (\underline{b} \mathbf{x} \underline{c}) = (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \bullet \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

From the basic formula for scalar product, this becomes

$$\underline{a} \bullet (\underline{b} \mathbf{x} \underline{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c3 \end{vmatrix}.$$

Notes:

(i) From Unit 7.3, if two rows of a determinant are interchanged, the determinant remains unchanged in numerical value but is altered in sign.

Hence,

$$\underline{a} \bullet (\underline{b} \mathbf{x} \underline{c}) = -\underline{a} \bullet (\underline{c} \mathbf{x} \underline{b}) = \underline{c} \bullet (\underline{a} \mathbf{x} \underline{b}) = -\underline{c} \bullet (\underline{b} \mathbf{x} \underline{a}) = \underline{b} \bullet (\underline{c} \mathbf{x} \underline{a}) = -\underline{b} \bullet (\underline{a} \mathbf{x} \underline{c}).$$

In other words, the **"cyclic permutations"** of $\underline{a} \bullet (\underline{b} \mathbf{x} \underline{c})$ are all equal in numerical value and in sign, while the remaining permutations are equal to $\underline{a} \bullet (\underline{b} \mathbf{x} \underline{c})$ in numerical value, but opposite in sign.

(ii) The triple scalar product, $\underline{a} \bullet (\underline{b} \mathbf{x} \underline{c})$, is often denoted by $[\underline{a}, \underline{b}, \underline{c}]$.

EXAMPLE

Evaluate the triple scalar product, $\underline{a} \bullet (\underline{b} \mathbf{x} \underline{c})$, in the case when

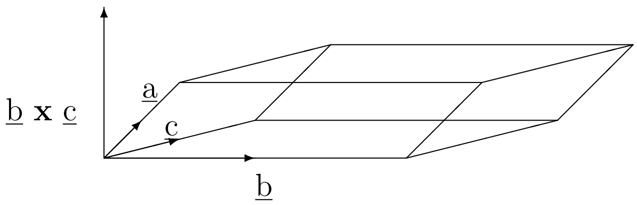
$$\underline{a} = 2\mathbf{i} + \mathbf{k}, \quad \underline{b} = \mathbf{i} + \mathbf{j} + 2\mathbf{k} \quad \text{and} \quad \underline{c} = -\mathbf{i} + \mathbf{j}$$

Solution

$$\underline{a} \bullet (\underline{b} \mathbf{x} \underline{c}) = \begin{vmatrix} 2 & 0 & 1 \\ 1 & 1 & 2 \\ -1 & 1 & 0 \end{vmatrix} = 2.(-2) - 0.(2) + 1.(2) = -2.$$

A geometrical application of the triple scalar product

Suppose that the three vectors $\underline{a}$, $\underline{b}$ and $\underline{c}$ lie along three adjacent edges of a parallelepiped (correct pronunciation, *"parallel-epi-ped"*) as shown in the following diagram:



The area of the base of the parallelepiped, from the geometrical properties of vector products, is the **magnitude** of the vector, $\underline{b}\ \mathbf{x}\ \underline{c}$, which is perpendicular to the base.

The perpendicular height of the parallelepiped is the projection of the vector $\underline{a}$ onto the vector $\underline{b}\ \mathbf{x}\ \underline{c}$; that is,

$$\frac{\underline{a}\ \bullet\ (\underline{b}\ \mathbf{x}\ \underline{c})}{|\underline{b}\ \mathbf{x}\ \underline{c}|}.$$

Hence, since the volume, V, of the parallelepiped is equal to the area of the base times the perpendicular height, we conclude that

$$V = \underline{a}\ \bullet\ (\underline{b}\ \mathbf{x}\ \underline{c}),$$

at least numerically, since the triple scalar product could turn out to be negative.

Note:
The above geometrical application also provides a condition that three given vectors, $\underline{a}$, $\underline{b}$ and $\underline{c}$ lie in the same plane; that is, they are **"coplanar"**.

The condition is that

$$\underline{a}\ \bullet\ (\underline{b}\ \mathbf{x}\ \underline{c}) = 0,$$

since the three vectors would determine a parallelepiped whose volume is zero.

8.4.2 THE TRIPLE VECTOR PRODUCT

DEFINITION 2

If $\underline{a}$, $\underline{b}$ and $\underline{c}$ are any three vectors, then the expression

$$\underline{a} \mathbf{\ x\ } (\underline{b} \mathbf{\ x\ } \underline{c})$$

is called the **"triple vector product"** of $\underline{a}$ with $\underline{b}$ and $\underline{c}$.

Notes:

(i) The triple vector product is clearly a vector quantity.

(ii) The inclusion of the brackets in a triple vector product is important since it can be shown that, in general,

$$\underline{a} \mathbf{\ x\ } (\underline{b} \mathbf{\ x\ } \underline{c}) \neq (\underline{a} \mathbf{\ x\ } \underline{b}) \mathbf{\ x\ } \underline{c}.$$

For example, if the three vectors are considered as position vectors, with the origin as a common end-point, then $\underline{a} \mathbf{\ x\ } (\underline{b} \mathbf{\ x\ } \underline{c})$ is perpendicular to both $\underline{a}$ and $\underline{b} \mathbf{\ x\ } \underline{c}$, the latter of which is already perpendicular to both $\underline{b}$ and $\underline{c}$. It therefore lies in the plane of $\underline{b}$ and $\underline{c}$.

Consequently, $(\underline{a} \mathbf{\ x\ } \underline{b}) \mathbf{\ x\ } \underline{c}$, which is the same as $-\underline{c} \mathbf{\ x\ } (\underline{a} \mathbf{\ x\ } \underline{b})$, will lie in the plane of $\underline{a}$ and $\underline{b}$.

Hence it will, in general, be different from $\underline{a} \mathbf{\ x\ } (\underline{b} \mathbf{\ x\ } \underline{c})$.

The formula for a triple vector product

Suppose that

$$\underline{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}, \quad \underline{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k} \ \text{ and } \ \underline{c} = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}.$$

Then,

$$\underline{a} \mathbf{\ x\ } (\underline{b} \mathbf{\ x\ } \underline{c}) = \underline{a} \mathbf{\ x\ } \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ (b_2c_3 - b_3c_2) & (b_3c_1 - b_1c_3) & (b_1c_2 - b_2c_1) \end{vmatrix}.$$

The **i** component of this vector is equal to

$$a_2(b_1c_2 - b_2c_1) - a_3(b_3c_1 - b_1c_3) = b_1(a_2c_2 + a_3c_3) - c_1(a_2b_2 + a_3b_3);$$

but, by adding and subtracting $a_1b_1c_1$, the right hand side can be rearranged in the form

$$(a_1c_1 + a_2c_2 + a_3c_3)b_1 - (a_1b_1 + a_2b_2 + a_3b_3)c_1,$$

which is the **i** component of the vector $(\underline{\mathrm{a}} \bullet \underline{\mathrm{c}})\underline{\mathrm{b}} - (\underline{\mathrm{a}} \bullet \underline{\mathrm{b}})\underline{\mathrm{c}}$.

Similar expressions can be obtained for the **j** and **k** components and we may conclude that

$$\underline{\mathrm{a}}\ \mathbf{x}\ (\underline{\mathrm{b}}\ \mathbf{x}\ \underline{\mathrm{c}}) = (\underline{\mathrm{a}} \bullet \underline{\mathrm{c}})\underline{\mathrm{b}} - (\underline{\mathrm{a}} \bullet \underline{\mathrm{b}})\underline{\mathrm{c}}.$$

EXAMPLE

Determine the triple vector product of $\underline{\mathrm{a}}$ with $\underline{\mathrm{b}}$ and $\underline{\mathrm{c}}$, where

$$\underline{\mathrm{a}} = \mathbf{i} + 2\mathbf{j} - \mathbf{k}, \quad \underline{\mathrm{b}} = -2\mathbf{i} + 3\mathbf{j} \quad \text{and} \quad \underline{\mathrm{c}} = 3\mathbf{k}.$$

Solution

$$\underline{\mathrm{a}} \bullet \underline{\mathrm{c}} = -3 \quad \text{and} \quad \underline{a} \bullet \underline{\mathrm{b}} = 4.$$

Hence,

$$\underline{\mathrm{a}}\ \mathbf{x}\ (\underline{\mathrm{b}}\ \mathbf{x}\ \underline{\mathrm{c}}) = -3\underline{\mathrm{b}} - 4\underline{\mathrm{c}} = 6\mathbf{i} - 9\mathbf{j} - 12\mathbf{k}.$$

8.4.3 EXERCISES

1. Evaluate the triple scalar product, $\underline{a} \bullet (\underline{b} \ \mathbf{x} \ \underline{c})$, in the case when

$$\underline{a} = 2\mathbf{i} + \mathbf{j} - 3\mathbf{k}, \quad \underline{b} = \mathbf{i} - 2\mathbf{j} + \mathbf{k} \quad \text{and} \quad \underline{c} = -\mathbf{i} + \mathbf{j} - 4\mathbf{k}.$$

2. Determine the volume of the parallelepiped with adjacent edges defined by the vectors

$$\underline{a} = -\mathbf{i} + \mathbf{j} + 3\mathbf{k}, \quad \underline{b} = 2\mathbf{i} - \mathbf{j} \quad \text{and} \quad \underline{c} = \mathbf{i} + \mathbf{j} + \mathbf{k}.$$

3. Determine the triple vector product of $\underline{a}$ with $\underline{b}$ and $\underline{c}$ in the cases where

   (a)

$$\underline{a} = -\mathbf{i} + \mathbf{j} + 3\mathbf{k}, \quad \underline{b} = 2\mathbf{i} + \mathbf{j} \quad \text{and} \quad \underline{c} = \mathbf{i} + \mathbf{j} + \mathbf{k};$$

   (b)

$$\underline{a} = 4\mathbf{i} - \mathbf{k}, \quad \underline{b} = 3\mathbf{i} + 5\mathbf{j} - \mathbf{k} \quad \text{and} \quad \underline{c} = \mathbf{i} - \mathbf{j} - \mathbf{k};$$

   (c)

$$\underline{a} = 4\mathbf{i} + 2\mathbf{j} + \mathbf{k}, \quad \underline{b} = 5\mathbf{i} \quad \text{and} \quad \underline{c} = -\mathbf{j} + 3\mathbf{k}.$$

4. Show that the following three vectors are coplanar:

$$\underline{a} = -\mathbf{i} + 3\mathbf{j} - 2\mathbf{k}, \quad \underline{b} = 4\mathbf{i} - 6\mathbf{j} + 2\mathbf{k} \quad \text{and} \quad \underline{c} = -3\mathbf{i} + 12\mathbf{j} - 9\mathbf{k}.$$

5. Show that

$$[(\underline{a} + \underline{b}), (\underline{b} + \underline{c}), (\underline{c} + \underline{a})] = 2[\underline{a}, \underline{b}, \underline{c}].$$

6. Show that

$$(\underline{a} \ \mathbf{x} \ \underline{b}) \ \mathbf{x}(\underline{c} \ \mathbf{x} \ \underline{d}) = [\underline{a}, \underline{c}, \underline{d}]\underline{b} - [\underline{b}, \underline{c}, \underline{d}]\underline{a} = [\underline{a}, \underline{b}, \underline{d}]\underline{c} - [\underline{a}, \underline{b}, \underline{c}]\underline{d}.$$

7. Show that

$$\underline{a} \ \mathbf{x} \ (\underline{b} \ \mathbf{x} \ \underline{c}) + \underline{b} \ \mathbf{x} \ (\underline{c} \ \mathbf{x} \ \underline{a}) + \underline{c} \ \mathbf{x} \ (\underline{a} \ \mathbf{x} \ \underline{b}) = \mathbf{O}.$$

8. Show that

$$(\underline{a} \ \mathbf{x} \ \underline{b}) \bullet (\underline{c} \ \mathbf{x} \ \underline{d}) = \begin{vmatrix} \underline{a} \bullet \underline{c} & \underline{a} \bullet \underline{d} \\ \underline{b} \bullet \underline{c} & \underline{b} \bullet \underline{d} \end{vmatrix}.$$

8.4.4 ANSWERS TO EXERCISES

1.
$$20.$$

2.
$$8.$$

3. (a)
$$7\mathbf{i} + 4\mathbf{j} + \mathbf{k};$$

   (b)
$$2\mathbf{i} + 38\mathbf{j} + 8\mathbf{k};$$

   (c)
$$5\mathbf{i} + 20\mathbf{j} - 60\mathbf{k}.$$

4. Show that the triple scalar product is zero.
5. Remove all brackets and use the fact that a triple scalar product is zero when two of the vectors are the same.
6. Use the triple vector product formula.
7. Use the triple vector product formula.
8. Rearrange in the form

$$\underline{a} \bullet [\underline{b} \mathbf{x} (\underline{c} \mathbf{x} \underline{d})].$$

"JUST THE MATHS"

UNIT NUMBER

8.5

VECTORS 5
(Vector equations of straight lines)

by

A.J.Hobson

UNIT 8.5 - VECTORS 5

VECTOR EQUATIONS OF STRAIGHT LINES

8.5.1 INTRODUCTION

The concept of vector notation and vector products provides a convenient method of representing straight lines and planes in space by simple vector equations. Such vector equations may then, if necessary, be converted back to conventional cartesian or parametric equations.

We shall assume that the position vector of a variable point, $\mathrm{P}(x, y, z)$, is given by

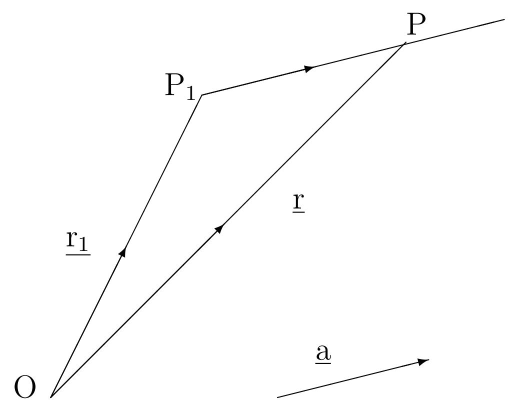
$$\underline{\mathrm{r}} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k},$$

and that the position vectors of fixed points, such as $\mathrm{P}_1(x_1, y_1, z_1)$ and $\mathrm{P}_2(x_2, y_2, z_2)$, are given by

$$\underline{\mathrm{r}_1} = x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k}, \quad \underline{\mathrm{r}_2} = x_2\mathbf{i} + y_2\mathbf{j} + z_2\mathbf{k}, \quad \text{etc.}$$

8.5.2 THE STRAIGHT LINE PASSING THROUGH A GIVEN POINT AND PARALLEL TO A GIVEN VECTOR

We consider, here, the straight line passing through the point, P_1, with position vector, $\underline{\mathrm{r}_1}$, and parallel to the vector, $\underline{\mathrm{a}} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$.



From the diagram,

$$\underline{\mathrm{OP}} = \underline{\mathrm{OP}_1} + \underline{\mathrm{P}_1\mathrm{P}}.$$

But,

$$\underline{\mathrm{P}_1\mathrm{P}} = t\underline{\mathrm{a}},$$

for some number t.

Hence,

$$\underline{\mathrm{r}} = \underline{\mathrm{r}_1} + t\underline{\mathrm{a}},$$

which is the vector equation of the straight line.

The components of $\underline{\mathrm{a}}$ form a set of direction ratios for the straight line.

Notes:

(i) The vector equation of a straight line passing through the origin and parallel to a given vector $\underline{\mathrm{a}}$ will be of the form

$$\underline{r} = t\underline{\mathrm{a}}.$$

(ii) By equating **i**, **j** and **k** components on both sides, the vector equation of the straight line passing through P_1 and parallel to $\underline{\mathrm{a}}$ leads to parametric equations

$$x = x_1 + a_1 t, \quad y = y_1 + a_2 t, \quad z = z_1 + a_3 t;$$

and, if these are solved for the parameter, t, we obtain

$$\frac{x - x_1}{a_1} = \frac{y - y_1}{a_2} = \frac{z - z_1}{a_3},$$

which is the standard cartesian form of the straight line.

EXAMPLES

1. Determine the vector equation of the straight line passing through the point with position vector $\mathbf{i} - 3\mathbf{j} + \mathbf{k}$ and parallel to the vector, $2\mathbf{i} + 3\mathbf{j} - 4\mathbf{k}$.

   Express the vector equation of the straight line in standard cartesian form.

   Solution

   The vector equation of the straight line is

   $$\underline{\mathrm{r}} = \mathbf{i} - 3\mathbf{j} + \mathbf{k} + t(2\mathbf{i} + 3\mathbf{j} - 4\mathbf{k})$$

   or

   $$x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = (1 + 2t)\mathbf{i} + (-3 + 3t)\mathbf{j} + (1 - 4t)\mathbf{k}.$$

   Eliminating t from each component, we obtain the cartesian form of the straight line,

   $$\frac{x-1}{2} = \frac{y+3}{3} = \frac{z-1}{-4}.$$

2. The equations

   $$\frac{3x+1}{2} = \frac{y-1}{2} = \frac{-z+5}{3}$$

   determine a straight line. Express them in vector form and obtain a set of direction ratios for the straight line.

   Solution

   Rewriting the equations so that the coefficients of x, y and z are unity, we have

   $$\frac{x+\frac{1}{3}}{\frac{2}{3}} = \frac{y-1}{2} = \frac{z-5}{-3}.$$

   Hence, in vector form, the equation of the line is

   $$\underline{\mathrm{r}} = -\frac{1}{3}\mathbf{i} + \mathbf{j} + 5\mathbf{k} + t\left(\frac{2}{3}\mathbf{i} + 2\mathbf{j} - 3\mathbf{k}\right).$$

   Thus, a set of direction ratios for the straight line are $\frac{2}{3} : 2 : -3$ or $2 : 6 : -9$.

3. Show that the two straight lines

$$\underline{\mathrm{r}} = \underline{\mathrm{r}_1} + t\underline{\mathrm{a}_1} \quad \text{and} \quad \underline{\mathrm{r}} = \underline{\mathrm{r}_2} + t\underline{\mathrm{a}_2},$$

where

$$\underline{\mathrm{r}_1} = \mathbf{j}, \quad \underline{\mathrm{a}_1} = \mathbf{i} + 2\mathbf{j} - \mathbf{k},$$

and

$$\underline{\mathrm{r}_2} = \mathbf{i} + \mathbf{j} + \mathbf{k}, \quad \underline{\mathrm{a}_2} = -2\mathbf{i} - 2\mathbf{j},$$

have a common point and determine its co-ordinates.

Solution

Any point on the first line is such that

$$x = t, \quad y = 1 + 2t, \quad z = -t,$$

for some parameter value, t; and any point on the second line is such that

$$x = 1 - 2s, \quad y = 1 - 2s, \quad z = 1,$$

for some parameter value, s.

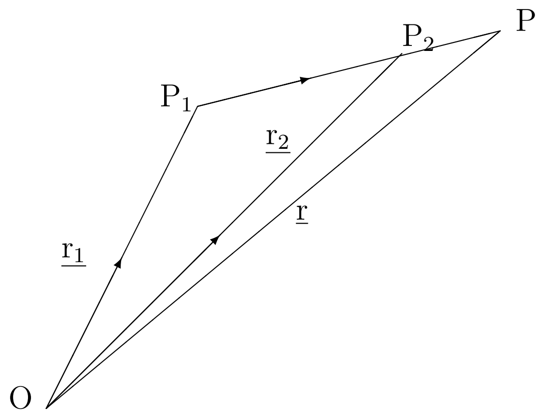
The lines have a common point if it is possible to find values of t and s such these are the same point.

In fact, $t = -1$ and $s = 1$ are suitable values and give the common point $(-1, -1, 1)$.

8.5.3 THE STRAIGHT LINE PASSING THROUGH TWO GIVEN POINTS

If a straight line passes through the two given points, P_1 and P_2, then it is certainly parallel to the vector,

$$\underline{\mathrm{a}} = \underline{\mathrm{P}_1\mathrm{P}_2} = (x_2 - x_1)\mathbf{i} + (y_2 - y_1)\mathbf{j} + (z_2 - z_1)\mathbf{k}.$$



Thus, the vector equation of the straight line is

$$\underline{\mathrm{r}} = \underline{\mathrm{r}_1} + t\underline{\mathrm{a}}$$

as before.

Notes:

(i) The parametric equations of the straight line passing through the points, P_1 and P_2, are

$$x = x_1 + (x_2 - x_1)t, \quad y = y_1 + (y_2 - y_1)t, \quad z = z_1 + (z_2 - z_1)t;$$

and we notice that the "base-points" of the parametric representation (that is, P_1 and P_2) have parameter values $t = 0$ and $t = 1$ respectively.

(ii) The standard cartesian form of the straight line passing through P_1 and P_2 is

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}.$$

1205

EXAMPLE

Determine the vector equation of the straight line passing through the two points, $P_1(3, -1, 5)$ and $P_2(-1, -4, 2)$.

Solution

$$\underline{OP_1} = 3\mathbf{i} - \mathbf{j} + 5\mathbf{k}$$

and

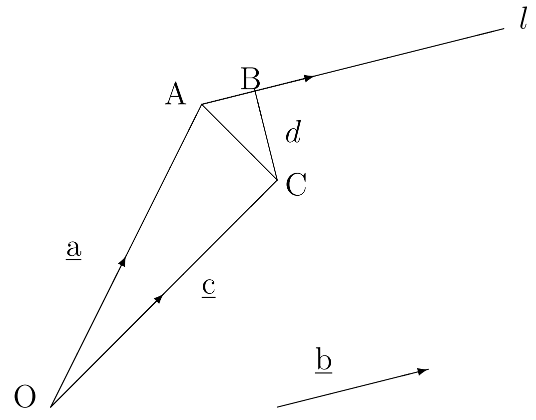
$$\underline{P_1P_2} = (-1-3)\mathbf{i} + (-4+1)\mathbf{j} + (2-5)\mathbf{k} = -4\mathbf{i} - 3\mathbf{j} - 3\mathbf{k}.$$

Hence, the vector equation of the straight line is

$$\underline{r} = 3\mathbf{i} - \mathbf{j} + 5\mathbf{k} - t(4\mathbf{i} + 3\mathbf{j} + 3\mathbf{k}).$$

8.5.4 THE PERPENDICULAR DISTANCE OF A POINT FROM A STRAIGHT LINE

For a straight line, l, passing through a given point, A, with position vector, $\underline{a}$ and parallel to a given vector, $\underline{b}$, it may be necessary to determine the perpendicular distance, d, from this line, of a point, C, with position vector, $\underline{c}$.



From the diagram, with Pythagoras' Theorem,

$$d^2 = (\mathrm{AC})^2 - (\mathrm{AB})^2.$$

But, $\underline{\mathrm{AC}} = \underline{\mathrm{c}} - \underline{\mathrm{a}}$, so that

$$(\mathrm{AC})^2 = (\underline{\mathrm{c}} - \underline{\mathrm{a}}) \bullet (\underline{\mathrm{c}} - \underline{\mathrm{a}}).$$

Also, the length, AB, is the projection of $\underline{\mathrm{AC}}$ onto the line, l, which is parallel to $\underline{\mathrm{b}}$.

Hence,

$$\mathrm{AB} = \frac{(\underline{\mathrm{c}} - \underline{\mathrm{a}}) \bullet \underline{\mathrm{b}}}{\mathrm{b}},$$

which gives the result

$$d^2 = (\underline{\mathrm{c}} - \underline{\mathrm{a}}) \bullet (\underline{\mathrm{c}} - \underline{\mathrm{a}}) - \left[\frac{(\underline{\mathrm{c}} - \underline{\mathrm{a}}) \bullet \underline{\mathrm{b}}}{\mathrm{b}}\right]^2.$$

From this result, d may be deduced.

EXAMPLE

Determine the perpendicular distance of the point $(3, -1, 7)$ from the straight line passing through the two points, $(2, 2, -1)$ and $(0, 3, 5)$.

Solution

In the standard formula, we have

$$\underline{\mathrm{a}} = 2\mathbf{i} + 2\mathbf{j} - \mathbf{j},$$

$$\underline{b} = (0-2)\mathbf{i} + (3-2)\mathbf{j} + (5-[-1])\mathbf{k} = -2\mathbf{i} + \mathbf{j} + 6\mathbf{k},$$

$$\mathrm{b} = \sqrt{(-2)^2 + 1^2 + 6^2} = \sqrt{41},$$

$$\underline{c} = 3\mathbf{i} - \mathbf{j} + 7\mathbf{k},$$

and

$$\underline{c} - \underline{a} = (3-2)\mathbf{i} + (-1-2)\mathbf{j} + (7-[-1])\mathbf{k} = \mathbf{i} - 3\mathbf{j} + 8\mathbf{k}.$$

Hence, the perpendicular distance, d, is given by

$$d^2 = 1^2 + (-3)^2 + 8^2 - \frac{(1)(-2)+(-3)(1)+(8)(6)}{\sqrt{41}} = 74 - \frac{43}{\sqrt{41}}$$

which gives $d \simeq 8.20$.

8.5.5 THE SHORTEST DISTANCE BETWEEN TWO PARALLEL STRAIGHT LINES

The result of the previous section may also be used to determine the shortest distance between two parallel straight lines, because this will be the perpendicular distance from one of the lines of any point on the other line.

We may consider the perpendicular distance between

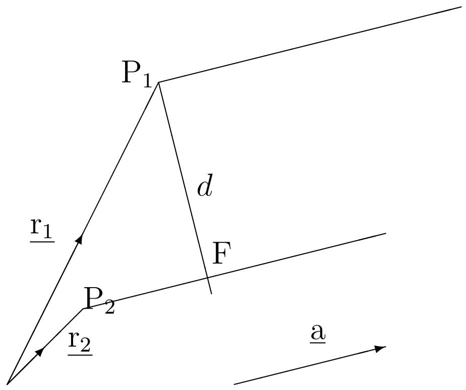
(a) the straight line passing through the fixed point with position vector $\underline{r_1}$ and parallel to the fixed vector, $\underline{a}$

and

(b) the straight line passing through the fixed point with position vector $\underline{r_2}$, also parallel to the fixed vector, $\underline{a}$.

These will have vector equations,

$$\underline{r} = \underline{r_1} + t\underline{a} \quad \text{and} \quad \underline{r} = \underline{r_2} + t\underline{a}.$$



In the diagram, F is the foot of the perpendicular onto the second line from the point P_1 on the first line and the length of this perpendicular is d.

Hence,

$$d^2 = (\underline{r_2} - \underline{r_1}) \bullet (\underline{r_2} - \underline{r_1}) - \left[\frac{(\underline{r_2} - \underline{r_1}) \bullet \underline{a}}{a}\right]^2.$$

EXAMPLE

Determine the shortest distance between the straight line passing through the point with position vector $\underline{r_1} = 4\mathbf{i} - \mathbf{j} + \mathbf{k}$, parallel to the vector $\underline{b} = \mathbf{i} + \mathbf{j} + \mathbf{k}$, and the straight line passing through the point with position vector $\underline{r_2} = -2\mathbf{i} + 3\mathbf{j} - \mathbf{k}$, parallel to $\underline{b}$.

Solution

From the formula,

$$d^2 = (-6\mathbf{i} + 4\mathbf{j} - 2\mathbf{k}) \bullet (-6\mathbf{i} + 4\mathbf{j} - 2\mathbf{k}) - \left[\frac{(-6\mathbf{i} + 4\mathbf{j} - 2\mathbf{k}) \bullet (\mathbf{i} + \mathbf{j} + \mathbf{k})}{\sqrt{3}}\right]^2.$$

That is,

$$d^2 = (36 + 16 + 4) - \left[\frac{-6 + 4 - 2}{\sqrt{3}}\right]^2 = 56 - \frac{16}{3} = \frac{152}{3},$$

which gives

$$d \simeq 7.12$$

8.5.6 THE SHORTEST DISTANCE BETWEEN TWO SKEW STRAIGHT LINES

Two straight lines are said to be **"skew"** if they are not parallel and do not intersect each other. It may be shown that such a pair of lines will always have a common perpendicular (that is, a straight line segment which meets both, and is perpendicular to both). Its length will be the shortest distance between the two skew lines.

For the straight lines, whose vector equations are

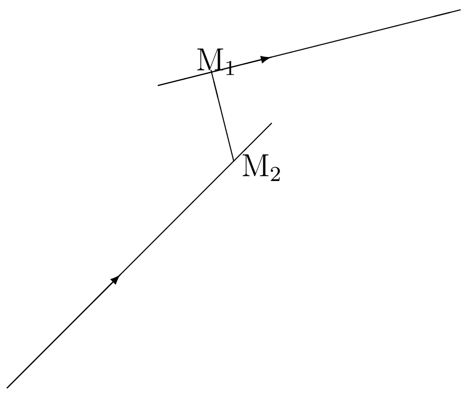
$$\underline{\mathrm{r}} = \underline{\mathrm{r}_1} + t\underline{\mathrm{a}_1} \ \text{ and } \ \underline{\mathrm{r}} = \underline{\mathrm{r}_2} + t\underline{\mathrm{a}_2},$$

suppose that the point, M_1, on the first line and the point, M_2, on the second line are the ends of the common perpendicular and have position vectors, $\underline{\mathrm{m}_1}$ and $\underline{\mathrm{m}_2}$, respectively.

Then,

$$\underline{\mathrm{m}_1} = \underline{\mathrm{r}_1} + t_1\underline{\mathrm{a}_1} \ \text{ and } \ \underline{\mathrm{m}_2} = \underline{\mathrm{r}_2} + t_2\underline{\mathrm{a}_2},$$

for some values, t_1 and t_2, of the parameter, t.



Firstly, we have

$$\underline{M_1M_2} = \underline{m_2} - \underline{m_1} = (\underline{r_2} - \underline{r_1}) + t_2\underline{a_2} - t_1\underline{a_1}.$$

Secondly, a vector which is certainly perpendicular to both of the skew lines is $\underline{a_1}$ $\mathbf{x}$ $\underline{a_2}$, so that a <u>unit</u> vector perpendicular to both of the skew lines is

$$\frac{\underline{a_1}\ \mathbf{x}\ \underline{a_2}}{|\underline{a_1}\ \mathbf{x}\ \underline{a_2}|}.$$

This implies that

$$(\underline{r_2} - \underline{r_1}) + t_2\underline{a_2} - t_1\underline{a_1} = \pm d\frac{\underline{a_1}\ \mathbf{x}\ \underline{a_2}}{|\underline{a_1}\ \mathbf{x}\ \underline{a_2}|},$$

where d is the shortest distance between the skew lines.

Finally, if we take the scalar (dot) product of both sides of this result with the vector $\underline{a_1}$ $\mathbf{x}$ $\underline{a_2}$, we obtain

$$(\underline{r_2} - \underline{r_1}) \bullet (\underline{a_1}\ \mathbf{x}\ \underline{a_2}) = \pm d\frac{|\underline{a_1}\ \mathbf{x}\ \underline{a_2}|^2}{|\underline{a_1}\ \mathbf{x}\ \underline{a_2}|},$$

giving

$$d = \left| \frac{(\underline{r_2} - \underline{r_1}) \bullet (\underline{a_1}\ \mathbf{x}\ \underline{a_2})}{|\underline{a_1}\ \mathbf{x}\ \underline{a_2}|} \right|.$$

EXAMPLE

Determine the perpendicular distance between the two skew lines

$$\underline{r} = \underline{r_1} + t\underline{a_1} \quad \text{and} \quad \underline{r} = \underline{r_2} + t\underline{a_2},$$

where

$$\underline{r_1} = 9\mathbf{j} + 2\mathbf{k}, \quad \underline{a_1} = 3\mathbf{i} - \mathbf{j} + \mathbf{k},$$

$$\underline{r}_2 = -6\mathbf{i} - 5\mathbf{j} + 10\mathbf{k}, \quad \underline{a}_2 = -3\mathbf{i} + 2\mathbf{j} + 4\mathbf{k}.$$

Solution

$$\underline{r}_2 - \underline{r}_1 = -6\mathbf{i} - 14\mathbf{j} + 8\mathbf{k}$$

and

$$\underline{a}_1 \ \mathbf{x} \ \underline{a}_2 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & -1 & 1 \\ -3 & 2 & 4 \end{vmatrix} = -6\mathbf{i} - 15\mathbf{j} + 3\mathbf{k},$$

so that

$$d = \frac{(-6)(-6) + (-14)(-15) + (8)(3)}{\sqrt{36 + 225 + 9}} = \frac{270}{\sqrt{270}} = \sqrt{270} = 3\sqrt{30}.$$

8.5.7 EXERCISES

1. Determine the vector equation, and hence the parametric equations, of the straight line which passes through the point, $(5, -2, 1)$, and is parallel to the vector $\mathbf{i} - 3\mathbf{j} + 2\mathbf{k}$.

2. The equations

$$\frac{-x + 2}{7} = \frac{3y - 1}{5} = \frac{2z + 1}{3}$$

determine a straight line. Determine the equation of the line in vector form, and state a set of direction ratios for this line.

3. Show that there is a point common to the two straight lines

$$\underline{r} = \underline{r}_1 + t\underline{a}_1 \quad \text{and} \quad \underline{r} = \underline{r}_2 + t\underline{a}_2,$$

where

$$\underline{r}_1 = 3\mathbf{j} + 2\mathbf{k}, \quad \underline{r}_2 = -2\mathbf{i} + 2\mathbf{j} + 3\mathbf{k},$$

and

$$\underline{a_1} = 2\mathbf{i} - 4\mathbf{j} - 2\mathbf{k}, \quad \underline{a_2} = 9\mathbf{i} - 3\mathbf{j} - 6\mathbf{k}$$

Determine the co-ordinates of the common point.

4. Determine, in standard cartesian form, the equation of the straight line passing through the two points, $(-2, 4, 9)$ and $(2, -1, 6)$.
5. Determine the perpendicular distance of the point $(0, -2, 5)$ from the straight line which passes through the point $(1, -1, 3)$ and is parallel to the vector $3\mathbf{i} + \mathbf{j} + \mathbf{k}$.
6. Determine the shortest distance between the two parallel straight lines

$$\underline{r} = \underline{r_1} + t\underline{a} \quad \text{and} \quad \underline{r} = \underline{r_2} + t\underline{a},$$

where

$$\underline{r_1} = -3\mathbf{i} + 2\mathbf{j} + \mathbf{k}, \quad \underline{r_2} = 2\mathbf{i} - \mathbf{j} - \mathbf{k}$$

and

$$\underline{a} = \mathbf{i} + 5\mathbf{j} + \mathbf{k}.$$

7. Determine the shortest distance between the two skew straight lines

$$\underline{r} = \underline{r_1} + t\underline{a_1} \quad \text{and} \quad \underline{r} = \underline{r_2} + t\underline{a_2},$$

where

$$\underline{r_1} = \mathbf{i} - 2\mathbf{j} + \mathbf{k}, \quad \underline{r_2} = -\mathbf{i} + 3\mathbf{j} + 2\mathbf{k},$$

and

$$\underline{a_1} = \mathbf{i} - \mathbf{j} - \mathbf{k}, \quad \underline{a_2} = 5\mathbf{i} + \mathbf{j} - 2\mathbf{k}$$

8.5.8 ANSWERS TO EXERCISES

1.

$$\underline{r} = (5\mathbf{i} - 2\mathbf{j} + \mathbf{k} + t(\mathbf{i} - 3\mathbf{j} + \mathbf{k}),$$

giving

$$x = 5 + t, \quad y = -2 - 3t, \quad z = 1 + t.$$

2. In vector form, the equation of the line is

$$\underline{\mathrm{r}} = 2\mathbf{i} + \frac{1}{3}\mathbf{j} - \frac{1}{2}\mathbf{k} + t\left(-7\mathbf{i} + \frac{5}{3}\mathbf{j} - \frac{3}{2}\mathbf{k}\right),$$

and set of direction ratios is

$$-42 : 10 : 9$$

3. The common point has co-ordinates $(1, 1, 1)$.

4.

$$\frac{x+2}{4} = \frac{y-4}{-5} = \frac{z-9}{-3}.$$

5.

$$d = \sqrt{\frac{62}{11}} \simeq 2.37$$

6.

$$d = \sqrt{\frac{98}{3}} \simeq 5.72$$

7.

$$d = \frac{5\sqrt{6}}{6} \simeq 2.04$$

"JUST THE MATHS"

UNIT NUMBER

8.6

VECTORS 6
(Vector equations of planes)

by

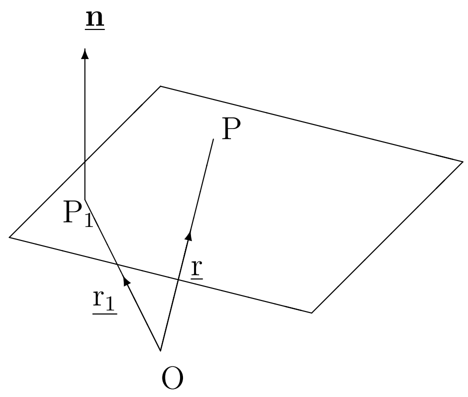
A.J.Hobson

UNIT 8.6 - VECTORS 6

VECTOR EQUATIONS OF PLANES

8.6.1 THE PLANE PASSING THROUGH A GIVEN POINT AND PERPENDICULAR TO A GIVEN VECTOR

A plane in space is completely specified if we know one point in it, together with a vector which is perpendicular to the plane.



In the diagram, the given point is P_1, with position vector, r_1, and the given vector is $\underline{n}$.

Hence, the vector, $\underline{P_1P}$, is perpendicular to $\underline{n}$, which leads to the equation

$$(\underline{r} - \underline{r_1}) \bullet \underline{n} = 0$$

or

$$\underline{r} \bullet \underline{n} = \underline{r_1} \bullet \underline{n} = d \text{ say.}$$

Notes:

(i) In the particular case when $\underline{n}$ is a unit vector, the constant, d, represents the perpendicular projection of $\underline{r_1}$ onto $\underline{n}$, which is therefore the perpendicular distance of the origin from the plane.

(ii) If $\underline{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ and $\underline{n} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$, then the cartesian form for the equation of the above plane will be

$$ax + by + cz = d.$$

That is, it is simply a linear equation in the variables x, y and z.

EXAMPLE

Determine the vector equation and, hence, the cartesian equation of the plane passing through the point with position vector $3\mathbf{i} - 2\mathbf{j} + \mathbf{k}$ and perpendicular to the vector $\mathbf{i} - 4\mathbf{j} - \mathbf{k}$.

Solution

The vector equation is

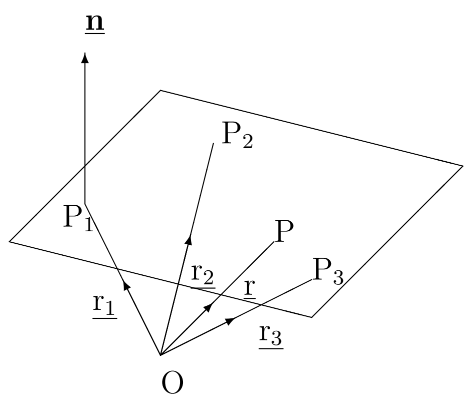
$$\underline{\mathrm{r}} \bullet (\mathbf{i} - 4\mathbf{j} - \mathbf{k}) = (3)(1) + (-2)(-4) + (1)(-1) = 10$$

and, hence, the cartesian equation is

$$x - 4y - z = 10.$$

8.6.2 THE PLANE PASSING THROUGH THREE GIVEN POINTS

We consider a plane passing through the points, $\mathrm{P}_1(x_1, y_1, z_1)$, $\mathrm{P}_2(x_2, y_2, z_2)$ and $\mathrm{P}_3(x_3, y_3, z_3)$.



In the diagram, a suitable vector for $\underline{\mathrm{n}}$ is

$$\underline{\mathrm{P_1P_2}}\ \mathbf{x}\ \underline{\mathrm{P_1P_3}} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix}$$

and, hence, the equation,

$$(\underline{\mathrm{r}} - \underline{\mathrm{r}_1}) \bullet \underline{\mathrm{n}} = 0,$$

of the plane becomes

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0.$$

But, from the properties of determinants, this is equivalent to

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0,$$

which is the standard equation of the plane through the three given points.

EXAMPLE

Determine the cartesian equation of the plane passing through the three points, $(0, 2, -1)$, $(3, 0, 1)$ and $(-3, -2, 0)$.

Solution

The equation of the plane is

$$\begin{vmatrix} x & y & z & 1 \\ 0 & 2 & -1 & 1 \\ 3 & 0 & 1 & 1 \\ -3 & -2 & 0 & 1 \end{vmatrix} = 0,$$

which, on expansion and simplification, gives

$$2x - 3y - 6z = 0,$$

showing that the plane also passes through the origin.

8.6.3 THE POINT OF INTERSECTION OF A STRAIGHT LINE AND A PLANE

First, we recall (from Unit 8.5) that the vector equation of a straight line passing through the fixed point, with position vector $\underline{\mathrm{r}}_1$, and parallel to the fixed vector $\underline{\mathrm{a}}$, is

$$\underline{\mathrm{r}} = \underline{\mathrm{r}_1} + t\underline{\mathrm{a}}.$$

For the point of intersection of this line with the plane, whose vector equation is

$$\underline{\mathrm{r}} \bullet \underline{\mathrm{n}} = d,$$

the value of t must be such that

$$(\underline{\mathrm{r}_1} + t\underline{\mathrm{a}}) \bullet \underline{\mathrm{n}} = d,$$

which is an equation from which the appropriate value of t and, hence, the point of intersection may be found.

EXAMPLE

Determine the point of intersection of the plane, whose vector equation is

$$\underline{\mathrm{r}} \bullet (\mathbf{i} - 3\mathbf{j} - \mathbf{k}) = 7,$$

and the straight line passing through the point, $(4, -1, 3)$, which is parallel to the vector $2\mathbf{i} - 2\mathbf{j} + 5\mathbf{k}$.

Solution

We need to obtain the value of the parameter, t, such that

$$(4\mathbf{i} - \mathbf{j} + 3\mathbf{k} + t[2\mathbf{i} - 2\mathbf{j} + 5\mathbf{k}]) \bullet (\mathbf{i} - 3\mathbf{j} - \mathbf{k}) = 7.$$

That is,

$$(4+2t)(1)+(-1-2t)(-3)+(3+5t)(-1)=7 \quad \text{or} \quad 4+3t=7,$$

which gives $t=1$; and, hence, the point of intersection is $(4+2,-1-2,3+5)=(6,-3,8)$.

8.6.4 THE LINE OF INTERSECTION OF TWO PLANES

Suppose we are given two non-parallel planes whose vector equations are

$$\underline{\mathrm{r}} \bullet \underline{\mathrm{n}_1} = d_1 \quad \text{and} \quad \underline{\mathrm{r}} \bullet \underline{\mathrm{n}_2} = d_2.$$

Their line of intersection will be perpendicular to both $\underline{\mathrm{n}_1}$ and $\underline{\mathrm{n}_2}$, since these are the normals to the two planes.

The line of intersection will thus be parallel to $\underline{\mathrm{n}_1}$ **x** $\underline{\mathrm{n}_2}$, and all it remains to do, to obtain the vector equation of this line, is to determine any point on it.

For convenience, we may take the point (common to both planes) for which one of x, y or z is zero.

EXAMPLE

Determine the vector equation, and hence the cartesian equations (in standard form), of the line of intersection of the planes whose vector equations are

$$\underline{\mathrm{r}} \bullet \underline{\mathrm{n}_1} = 2 \quad \text{and} \quad \underline{\mathrm{r}} \bullet \underline{\mathrm{n}_2} = 17,$$

where

$$\underline{\mathrm{n}_1} = \mathbf{i}+\mathbf{j}+\mathbf{k} \quad \text{and} \quad \underline{\mathrm{n}_2} = 4\mathbf{i}+\mathbf{j}+2\mathbf{k}.$$

Solution

Firstly,

$$\underline{\mathrm{n}_1} \ \mathbf{x} \ \underline{\mathrm{n}_2} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 1 & 1 \\ 4 & 1 & 2 \end{vmatrix} = \mathbf{i}+2\mathbf{j}-3\mathbf{k}.$$

Secondly, the cartesian equations of the two planes are

$$x + y + z = 2 \text{ and } 4x + y + 2z = 17;$$

and, when $z = 0$, these become

$$x + y = 2 \text{ and } 4x + y = 17,$$

which, as simultaneous linear equations, have a common solution of $x = 5$, $y = -3$.

Thirdly, therefore, a point on the line of intersection is $(5, -3, 0)$, which has position vector $5\mathbf{i} - 3\mathbf{j}$.

Hence, the vector equation of the line of intersection is

$$\underline{\mathrm{r}} = 5\mathbf{i} - 3\mathbf{j} + t(\mathbf{i} + 2\mathbf{j} - 3\mathbf{k}).$$

Finally, since $x = 5 + t$, $y = -3 + 2t$ and $z = -3t$ the line of intersection is represented, in standard cartesian form, by

$$\frac{x-5}{1} = \frac{y+3}{2} = \frac{z}{-3} \quad (= t).$$

8.6.5 THE PERPENDICULAR DISTANCE OF A POINT FROM A PLANE

Given the plane whose vector equation is $\underline{\mathrm{r}} \bullet \underline{\mathrm{n}} = d$ and the point, P_1, whose position vector is $\underline{\mathrm{r}_1}$, the straight line through the point, P_1, which is perpendicular to the plane has vector equation

$$\underline{\mathrm{r}} = \underline{\mathrm{r}_1} + t\underline{\mathrm{n}}.$$

This line meets the plane at the point, P_0, with position vector $\underline{\mathrm{r}_1} + t_0\underline{\mathrm{n}}$, where

$$(\underline{\mathrm{r}_1} + t_0\underline{\mathrm{n}}) \bullet \underline{\mathrm{n}} = d.$$

That is,

$$(\underline{\mathrm{r}_1} \bullet \underline{\mathrm{n}}) + t_0 \mathrm{n}^2 = d.$$

Hence,

$$t_0 = \frac{d - (\underline{\mathrm{r}_1} \bullet \underline{\mathrm{n}})}{\mathrm{n}^2}$$

Finally, the vector $\underline{\mathrm{P}_0\mathrm{P}_1} = (\underline{\mathrm{r}_1} + t_0\underline{\mathrm{n}}) - \underline{\mathrm{r}_1} = t_0\underline{\mathrm{n}}$

and its magnitude, $t_0\mathrm{n}$, will be the perpendicular distance, p, of the point P_1 from the plane.

In other words,

$$p = \frac{d - (\underline{\mathrm{r}_1} \bullet \underline{\mathrm{n}})}{\mathrm{n}}.$$

Note:
In terms of cartesian co-ordinates, this formula is equivalent to

$$p = \frac{d - (ax_1 + by_1 + cz_1)}{\sqrt{a^2 + b^2 + c^2}},$$

where a, b and c are the **i**, **j** and **k** components of $\underline{\mathrm{n}}$ respectively.

EXAMPLE

Determine the perpendicular distance, p, of the point $(2, -3, 4)$ from the plane whose cartesian equation is $x + 2y + 2z = 13$.

Solution

From the cartesian formula

$$p = \frac{13 - [(1)(2) + (2)(-3) + (2)(4)]}{\sqrt{1^2 + 2^2 + 2^2}} = \frac{9}{3} = 3.$$

8.6.6 EXERCISES

1. Determine the vector equation and hence the cartesian equation of the plane, passing through the point with position vector $\mathbf{i} + 5\mathbf{j} - \mathbf{k}$, and perpendicular to the vector $2\mathbf{i} + \mathbf{j} - 3\mathbf{k}$.

2. Determine the cartesian equation of the plane passing through the three points $(1, -1, 2)$, $(3, -2, -1)$ and $(-1, 4, 0)$.

3. Determine the point of intersection of the plane, whose vector equation is

$$\underline{\mathrm{r}} \bullet (5\mathbf{i} - 2\mathbf{j} - \mathbf{k}) = -3,$$

and the straight line passing through the point $(2, 1, -3)$, which is parallel to the vector $\mathbf{i} + \mathbf{j} - 4\mathbf{k}$.

4. Determine the vector equation, and hence the cartesian equations (in standard form), of the line of intersection of the planes, whose vector equations are

$$\underline{\mathrm{r}} \bullet \underline{\mathrm{n}_1} = 14 \quad \text{and} \quad \underline{\mathrm{r}} \bullet \underline{\mathrm{n}_2} = -1,$$

where

$$\underline{\mathrm{n}_1} = -4\mathbf{i} + 2\mathbf{j} - \mathbf{k} \quad \text{and} \quad \underline{\mathrm{n}_2} = 2\mathbf{i} + \mathbf{j} + 3\mathbf{k}.$$

5. Determine, in surd form, the perpendicular distance of the point $(-5, -2, 8)$ from the plane whose cartesian equation is $2x - y + 3z = 17$.

8.6.7 ANSWERS TO EXERCISES

1.

$$\underline{\mathrm{r}} \bullet (2\mathbf{i} + \mathbf{j} - 3\mathbf{k}) = 10, \quad \text{or} \quad 2x + y - 3z = 10.$$

2.

$$17x + 10y + 8z = 23.$$

3.

$$(0, -1, 5).$$

4.

$$\underline{\mathrm{r}} = -2\mathbf{i} + 3\mathbf{j} + t(7\mathbf{i} + 10\mathbf{j} - 8\mathbf{k})$$

or

$$\frac{x+2}{7} = \frac{y-3}{10} = \frac{z}{-8} \quad (= t).$$

5.

$$\frac{1}{\sqrt{14}}.$$

"JUST THE MATHS"

UNIT NUMBER

9.1

MATRICES 1
(Definitions & elementary matrix algebra)

by

A.J.Hobson

UNIT 9.1 - MATRICES 1
DEFINITIONS AND ELEMENTARY MATRIX ALGEBRA

9.1.1 INTRODUCTION

(a) Presentation of Data
Sets of numerical information can often be presented as a rectangular **"array"** of numbers. For example, a football results table might look like this:

TEAM	PLAYED	WON	DRAWN	LOST
Blackburn	22	11	6	5
Burnley	22	9	6	7
Chelsea	21	7	8	6
Leicester	21	6	8	7
Stoke	21	6	6	9

If the headings are taken for granted, we write simply

$$\begin{bmatrix} 22 & 11 & 6 & 5 \\ 22 & 9 & 6 & 7 \\ 21 & 7 & 8 & 6 \\ 21 & 6 & 8 & 7 \\ 21 & 6 & 6 & 9 \end{bmatrix}$$

and this symbol is called a **"matrix"**.

Note:
The mould in which printers once cast **type** was called a matrix. In mathematics, the word "matrix" signifies that we have spaces into which **numbers** can be placed.

(b) Presentation of Algebraic Results
In two-dimensional geometry, the **"vector"**, $\underline{\mathrm{OP}_0}$, joining the origin to the point $\mathrm{P}_0(x_0, y_0)$ can be moved to the position $\underline{\mathrm{OP}_1}$ joining the origin to the point $\mathrm{P}_1(x_1, y_1)$ by means of a **"reflection"**, a **"rotation"**, a **"magnification"** or a combination of such operations. It can be shown that the relationship between the co-ordinates of P_0 and P_1 in any such operation is given by

$$\begin{aligned} x_1 &= ax_0 + by_0, \\ y_1 &= cx_0 + dy_0. \end{aligned}$$

ILLUSTRATIONS

1. The equations

$$\begin{aligned} x_1 &= -x_0, \\ y_1 &= \quad\quad y_0 \end{aligned}$$

represent a reflection in the y-axis.

2. The equations

$$\begin{aligned} x_1 &= kx_0, \\ y_1 &= \quad\quad ky_0 \end{aligned}$$

represent a magnification when $| k |> 1$ and a contraction when $| k |< 1$.

3. The equations

$$\begin{aligned} x_1 &= x_0 \cos\theta - y_0 \sin\theta, \\ y_1 &= x_0 \sin\theta + y_0 \cos\theta \end{aligned}$$

represent a rotation of $\underline{\mathrm{OP}}_0$ through an angle θ in a counter-clockwise direction.

Such an operation is called a **"linear transformation"** and is completely specified by the coefficients a, b, c, d in the correct positions. When referring to a linear transformation, it is therefore more convenient to write the matrix symbol

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Sometimes, we want to show that a transformation

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

operates on the vector $\underline{\mathrm{OP}}_0$, transforming it into the vector $\underline{\mathrm{OP}}_1$. We write

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} . \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}.$$

The two new symbols are still called matrices but can be given a special name as we see later.

9.1.2 DEFINITIONS

(i) A matrix is **a rectangular array of numbers** arranged in rows (horizontally), columns (vertically) and enclosed in brackets.

ILLUSTRATIONS:

$$\begin{bmatrix} 1 & 3 & 7 \\ 9 & 9 & 10 \end{bmatrix}, \begin{bmatrix} 10 & 16 & 17 \\ 3 & 5 & 11 \\ 4 & 0 & 10 \end{bmatrix}, \begin{bmatrix} 1 & 5 & 7 \end{bmatrix}.$$

Note:
The rows are counted from the top to the bottom of the matrix and the columns are counted from the left to the right of the matrix.

(ii) Any number within the array is called an **"element"** of the matrix. The term $\underline{ij-th}$ element refers to the element lying in the i-th row and the j-th column of the matrix.

(iii) If a matrix has m rows and n columns, it is called a **"matrix of order** $m \times n$**"** or simply an **"**$m \times n$ **matrix"**. It clearly has mn elements.

(iv) A matrix of order $m \times m$ is called a **"square matrix"**.

Note:
A matrix of order 1×1 is considered to be the same as a single number. This assumption has particular significance in the definition of a matrix product (see later).

(v) A matrix of order $m \times 1$ is called a **"column vector"** and a matrix of order $1 \times n$ is called a **"row vector"**.

Note:
Matrices of order 2×1, 1×2, 3×1 and 1×3 have easy physical interpretations as vectors. But, in abstract linear algebra, there are vectors of **any** dimension; so the names "column vector" and "row vector" are retained no matter how many elements there are.

(vi) An arbitrary matrix whose elements and order do not have to be specified may be denoted by a single capital letter such as A,B,C, etc.

An arbitrary matrix of order $m \times n$ may be denoted fully by the symbol

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}$$

where, in each double-subscript, the first number is the row number and the second number is the column number. In a situation where the matrix has already been denoted by a single capital letter, the subsequent use of the full notation should include the corresponding small letter on which to attach the double-subscripts.

If the matrix is square, the elements $a_{11}, a_{22}, a_{33}, ...a_{mm}$ are called the **"diagonal elements"** and their sum is called the **"trace"** of the matrix. This line of elements is called the **"leading diagonal"** of the matrix.

An abbreviated form of the full notation is $[a_{ij}]_{m \times n}$

EXAMPLE
A matrix $\mathrm{A} = [a_{ij}]_{2\times 3}$ is such that $a_{ij} = i^2 + 2j$. Write out A in full.

Solution

This is an artificially contrived example, but will serve to illustrate the use of the double-subscript notation. We obtain

$$\mathrm{A} = \begin{bmatrix} 3 & 5 & 7 \\ 6 & 8 & 10 \end{bmatrix}.$$

(vii) Given a matrix, A, of order $m \times n$, the matrix of order $n \times m$ obtained from A by writing the rows as columns is called the **"transpose"** of A and is denoted by A^{T} - some books use A' or $\tilde{\mathrm{A}}$.

ILLUSTRATION

If

$$\mathrm{A} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix},$$

then,

$$\mathrm{A}^{\mathrm{T}} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}$$

(viii) A matrix, A, is said to be **"symmetric"** if $\mathrm{A} = \mathrm{A}^{\mathrm{T}}$.

ILLUSTRATION

$$\mathrm{A} = \begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix} \text{ is symmetric.}$$

(ix) A matrix, A, is said to be **"skew-symmetric"** if the elements of A^T are minus the corresponding elements of A itself. This will mean that the leading diagonal elements must be zero.

ILLUSTRATION

$$A = \begin{bmatrix} 0 & a & -b \\ -a & 0 & c \\ b & -c & 0 \end{bmatrix} \text{ is skew} - \text{symmetric.}$$

(x) A matrix is said to be **"diagonal"** if the elements which are not on the leading diagonal have value zero while the elements on the leading diagonal are not all equal to zero.

9.1.3 THE ALGEBRA OF MATRICES (Part One)

An **"Algebra"** (coming from the Arabic word AL-JABR) refers to any mathematical system which uses the concepts of equality, addition, subtraction, multiplication and division. For example, the algebra of numbers is what we normally call **"arithmetic"**; but algebraical concepts can be applied to other mathematical systems of which matrices is one.

In meeting a new mathematical system for the first time, the concepts of equality, addition, subtraction, multiplication and division need to be properly defined, and that is the purpose of the present section. In some cases, the definitions are fairly obvious, but they need to be made without contradicting ideas already established in the system of numbers which matrices depend on.

1. Equality

Unlike a determinant, a matrix does **not** have a numerical value; so the use of equality here is not that which is made in elementary arithmetic. Rather we use it to mean "is the same as".

Two matrices are said to be equal if they have **the same order** and also **pairs of elements in corresponding positions are equal in value**.

In symbols, we could say that

$$[a_{ij}]_{m\times n} = [b_{ij}]_{m\times n}$$

provided

$$a_{ij} = b_{ij}.$$

For example,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

implies that $x = 1, y = 2$ and $z = 3$.

We shall meet this type of conclusion later in a discussion on the solution, by matrices, of a set of simultaneous equations (see Unit 9.3).

2. Addition and Subtraction

The scientific applications of matrices require us to define the sum and difference of two matrices only when they have the same order.

The sum of two matrices of order $m \times n$ is formed by adding together the pairs of elements in corresponding positions. Similarly, the difference of two matrices of order $m \times n$ is formed by subtracting the pairs of elements in corresponding positions.

In symbols, we could say that

$$[a_{ij}]_{m\times n} \pm [b_{ij}]_{m\times n} = [c_{ij}]_{m\times n}\,,$$

where

$$c_{ij} = a_{ij} \pm b_{ij}.$$

Note:
It may occur, in a calculation, that a matrix has to be subtracted from itself, in which case we would obtain another matrix of the same order but whose elements are all zero. This type of matrix is called a **"null matrix"** and may be denoted, for short, by $[0]_{m\times n}$ or just $[0]$ when the order is understood.

ILLUSTRATION
A grocer has two shops and, in each shop, he sells apples, oranges and bananas. The sales, in kilogrammes, of each fruit for the two shops on two separate days are represented by the matrices

$$\begin{bmatrix} 36 & 25 & 10 \\ 20 & 30 & 15 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 40 & 30 & 12 \\ 22 & 35 & 20 \end{bmatrix}$$

where the rows refer to the two shops and the columns refer to apples, oranges and bananas respectively.

The total sales, in kilogrammes, of each fruit for the two shops on both days together are represented by the matrix

$$\begin{bmatrix} 76 & 55 & 22 \\ 42 & 65 & 35 \end{bmatrix}.$$

The differences in sales of each fruit for the two shops between the second day and the first day are represented by the matrix

$$\begin{bmatrix} 4 & 5 & 2 \\ 2 & 5 & 5 \end{bmatrix}$$

3. Additive Identities and Additive Inverses

For the sake of completeness, we mention here that

(a) A null matrix, when added to any other matrix of the same order, leaves that other matrix identically the same as it was to start with. For this reason, a null matrix behaves as an **"additive identity"**.

(b) If any matrix A is added to the matrix $-$A (that is, the matrix obtained from A by reversing the signs of all the elements) the result obtained is the corresponding null matrix. For this reason $-$A can be called the **"additive inverse"** of A.

9.1.4 EXERCISES

1. State the order of each of the following matrices:
   (a) $\begin{bmatrix} 1 & 3 & 0 \\ -1 & 2 & 4 \end{bmatrix}$; (b) $\begin{bmatrix} 5 \\ 6 \end{bmatrix}$; (c) $[3 \quad 5 \quad 4]$.
2. Two of the following matrices are equal to each other. Which are they ?
   (a) $\begin{bmatrix} \frac{1}{2} & \frac{1}{\sqrt{3}} \\ 1 & 0 \end{bmatrix}$; (b) $\begin{bmatrix} \sin 60^\circ & \tan 30^\circ \\ \tan 45^\circ & \sin 90^\circ \end{bmatrix}$; (c) $\begin{bmatrix} \cos 60^\circ & \frac{\sqrt{3}}{3} \\ \sin 90^\circ & \cos 90^\circ \end{bmatrix}$.
3. Determine the values of x, y and z given that
   $$\begin{bmatrix} x+y & y+z \\ x+z & 3x-2y \end{bmatrix} = \begin{bmatrix} 5 & 7 \\ 6 & 0 \end{bmatrix}.$$
4. If A $= \begin{bmatrix} 1 & -3 & 2 \\ 4 & 0 & 1 \end{bmatrix}$, B $= \begin{bmatrix} 4 & 3 & 2 \\ -1 & 0 & 2 \end{bmatrix}$ and C $= \begin{bmatrix} -1 & 4 & -2 \\ 0 & 1 & 0 \end{bmatrix}$,
   determine the elements of the following matrices:
   (a) A + B; (b) (A + B) + C; (c) A + (B + C);
   (d) A - B - C; (e) $\mathrm{A}^{\mathrm{T}} + \mathrm{B}^{\mathrm{T}}$; (f) $(\mathrm{A}+\mathrm{B})^{\mathrm{T}}$.

5. Determine which pairs of the following matrices can be added and, for those which can, state the sum:

$$A = \begin{bmatrix} 1 \\ 2 \end{bmatrix},\ B = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix},\ C = \begin{bmatrix} 4 & 2 & 1 \end{bmatrix},\ D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$E = \begin{bmatrix} 2 \\ 1 \end{bmatrix},\ F = \begin{bmatrix} 2 & 3 \\ 5 & 0 \end{bmatrix},\ G = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},\ H = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},\ J = \begin{bmatrix} 3 & 3 & 3 \end{bmatrix}.$$

9.1.5 ANSWERS TO EXERCISES

1. (a) 2×3; (b) 2×1; (c) 1×3.
2. (a) and (c) are equal to each other.
3. $x = 2, y = 3$ and $z = 4$.
4. (a) $\begin{bmatrix} 5 & 0 & 4 \\ 3 & 0 & 3 \end{bmatrix}$; (b) $\begin{bmatrix} 4 & 4 & 2 \\ 3 & 1 & 3 \end{bmatrix}$; (c) $\begin{bmatrix} 4 & 4 & 2 \\ 3 & 1 & 3 \end{bmatrix}$;

   (d) $\begin{bmatrix} -2 & -10 & 2 \\ 5 & -1 & -1 \end{bmatrix}$; (e) $\begin{bmatrix} 5 & 3 \\ 0 & 0 \\ 4 & 3 \end{bmatrix}$; (f) $\begin{bmatrix} 5 & 3 \\ 0 & 0 \\ 4 & 3 \end{bmatrix}$.
5. $A + E = E + A = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, $B + G = G + B = \begin{bmatrix} 5 \\ 2 \\ 2 \end{bmatrix}$

   $C + J = J + C = \begin{bmatrix} 7 & 5 & 4 \end{bmatrix}$, $D + F = F + D = \begin{bmatrix} 2 & 3 \\ 5 & 0 \end{bmatrix}$

   $G + H = H + G = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $B + H = H + B = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}$.

   That is, twelve possible additions but only six distinct results.

"JUST THE MATHS"

UNIT NUMBER

9.2

MATRICES 2
(Further matrix algebra)

by

A.J.Hobson

UNIT 9.2 - MATRICES 2 - THE ALGEBRA OF MATRICES (Part Two)

9.2.1 MULTIPLICATION BY A SINGLE NUMBER

If we were required to multiply a matrix of any order by a **positive whole number**, n, we would clearly regard the operation as equivalent to adding together n copies of the given matrix. Thus, in the result, every element of this given matrix would be multiplied by n; but it is logical to extend the idea to the multiplication of a matrix by **any** number, λ, not necessarily a positive whole number, the rule being to multiply every element of the matrix by λ.

In symbols we could say that

$$\lambda\,[a_{ij}]_{m\times n} = [b_{ij}]_{m\times n}\,,$$

where

$$b_{ij} = \lambda a_{ij}.$$

Note:
The rule for multiplying a matrix by a single number can also be used in reverse to remove common factors from the elements of a matrix as illustrated as follows:

ILLUSTRATION

$$\begin{bmatrix} 5 & 10 \\ 15 & 20 \end{bmatrix} = 5\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

9.2.2 THE PRODUCT OF TWO MATRICES

The definition of a matrix product is more difficult to justify than the previous concepts, partly because it is by no means an obvious definition and partly because we cannot be sure exactly what originally led to the making of the definition.

Some hint is given by the matrix equation at the end of the introduction to Unit 9.1, where the product of a 2×2 matrix and a 2×1 matrix is another 2×1 matrix; but we must be prepared to meet other orders of matrix as well.

We shall introduce the definition with a semi-practical illustration, then make a formal statement of the definition itself.

ILLUSTRATION

A motor manufacturer, with three separate factories, makes two types of car, one called "standard" and the other called "luxury".
In order to manufacture each type of car, he needs a certain number of units of material and a certain number of units of labour each unit representing £300.

A table of data to represent this information could be

Type	Materials	Labour
Standard	12	15
Luxury	16	20

The manufacturer receives an order from another country to supply 400 standard cars and 900 luxury cars; but he distributes the export order amongst his three factories as follows:

Location	Standard	Luxury
Factory A	100	400
Factory B	200	200
Factory C	100	300

The number of units of material and the number of units of labour needed by each factory to complete the order may be given by another table, namely

Location	Materials	Labour
Factory A	$100 \times 12 + 400 \times 16$	$100 \times 15 + 400 \times 20$
Factory B	$200 \times 12 + 200 \times 16$	$200 \times 15 + 200 \times 20$
Factory C	$100 \times 12 + 300 \times 16$	$100 \times 15 + 300 \times 20$

If we now replace each table by the corresponding matrix, the calculations appear as the product of a 3×2 matrix and a 2×2 matrix. That is,

$$\underset{3\times 2}{\begin{bmatrix} 100 & 400 \\ 200 & 200 \\ 100 & 300 \end{bmatrix}} \cdot \underset{2\times 2}{\begin{bmatrix} 12 & 15 \\ 16 & 20 \end{bmatrix}} = \underset{3\times 2}{\begin{bmatrix} 100 \times 12 + 400 \times 16 & 100 \times 15 + 400 \times 20 \\ 200 \times 12 + 200 \times 16 & 200 \times 15 + 200 \times 20 \\ 100 \times 12 + 300 \times 16 & 100 \times 15 + 300 \times 20 \end{bmatrix}} = \underset{3\times 2}{\begin{bmatrix} 7600 & 9500 \\ 5600 & 7000 \\ 6000 & 7500 \end{bmatrix}}$$

OBSERVATIONS

(i) The product matrix has 3 rows because the first matrix on the left has 3 rows.

(ii) The product matrix has 2 columns because the second matrix on the left has 2 columns.

(iii) The product cannot be worked out unless the number of columns in the first matrix matches the number of rows in the second matrix with no elements left over in the pairing-up process.

(iv) The elements of the product matrix are systematically obtained by multiplying (in pairs) the corresponding elements of each row in the first matrix with each column in the second matrix. To pair up the correct elements, we read each row of the first matrix from left to right and each column of the second matrix from top to bottom.

The Formal Definition of a Matrix Product
If A and B are matrices, then the product AB is defined (that is, it has a meaning) only when the number of columns in A is equal to the number of rows in B.

If A is of order $m \times n$ and B is of order $n \times p$, then AB is of order $m \times p$.

To obtain the element in the i-th row and j-th column of AB, we multiply corresponding elements of the i-th row of A and the j-th column of B, then add up the results.

ILLUSTRATION

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 1 & 6 \\ 3 & 2 & -1 & 1 \\ 1 & -2 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 8 & 2 & -1 & 12 \\ 1 & -10 & 1 & 15 \end{bmatrix}$$

Note:
Confusion could arise when multiplying a matrix of order 1×1 by another matrix. Apparently, the other matrix would need to have either a single row or a single column depending on the order of multiplication.

However, as stated in Unit 9.1, a matrix of order 1×1 is considered to be a special case, and is defined separately to be the same as a single number. Hence a matrix of any order can be multiplied by a matrix of order 1×1 even though this does not fit the formal rules for matrix multiplication in general.

9.2.3 THE NON-COMMUTATIVITY OF MATRIX PRODUCTS

In elementary arithmetic, if a and b are two numbers, then $ab = ba$ (that is, the product **"commutes"**). But this is not so for matrices A and B as we now show:

(a) If A is of order $m \times n$, then B must be of order $n \times m$ if both AB **and** BA are to be defined.

(b) AB and BA will have different orders unless $m = n$, in which case the two products will be square matrices of order $m \times m$.

(c) Even if A and B are **both** square matrices of order $m \times m$, it will not normally be the case that AB is the same as BA. A simple numerical example will illustrate this fact:

$$\begin{bmatrix} 2 & 1 \\ 3 & 7 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 0 & 5 \\ 11 & 35 \end{bmatrix}; \quad \text{but} \quad \begin{bmatrix} -1 & 0 \\ 2 & 5 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 \\ 3 & 7 \end{bmatrix} = \begin{bmatrix} -2 & -1 \\ 19 & 37 \end{bmatrix}.$$

Notes:
(i) If we simply wanted to show that $AB \neq BA$, we would need only to demonstrate that one pair of corresponding elements were unequal in value.

(ii) If such a basic rule of elementary arithmetic is false for matrices, we should, strictly speaking, be prepared to justify other basic rules of arithmetic. But it turns out that the non-commutativity of matrix products is the only one which causes problems.

For instance, it can be shown that, provided the matrices involved are compatible for addition or multiplication,

$A + B \equiv B + A$; the **"Commutative Law of Addition"**.

$A + (B + C) \equiv (A + B) + C$; the **"Associative Law of Addition"**.

$A(BC) \equiv (AB)C$; the **"Associative Law of Multiplication"**.

$A(B + C) \equiv AB + BC$ or $(A + B)C \equiv AC + BC$; the **"Distributive Laws"**.

(iii) In the matrix product, AB, we say either that B is **"pre-multiplied"** by A or that A is **"post-multiplied"** by B.

9.2.4 MULTIPLICATIVE IDENTITY MATRICES

In connection with matrix multiplication and, subsequently, the solution of simultaneous linear equations, an important type of matrix is a square matrix with the number 1 in each position of the leading diagonal, but zeros everywhere else. Such a matrix is denoted by I_n if there are n rows and (of course) n columns. If it is not necessary to specify the number of rows and columns, the notation I, without a subscript, is sufficient.

ILLUSTRATIONS

$$I_2 \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad I_3 \equiv \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad I_4 \equiv \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Any matrix of the type I_n multiplies another matrix (with an appropriate number of rows or columns) to leave it identically the same as it was to start with. For this reason, I_n is called a **"multiplicative identity matrix"**, although we normally call it just an "identity matrix" (unless it becomes necessary to distinguish it from the **additive** identity matrix referred to earlier). Another common name for it is a **"unit matrix"**.

For example, suppose that

$$A = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix}.$$

Then, post-multiplying by I_2, it is easily checked that

$$AI_2 = A.$$

Similarly, pre-multiplying by I_3, it is easily checked that

$$I_3A = A.$$

In general, if A is of order $m \times n$, then

$$AI_n = I_mA = A.$$

9.2.5 EXERCISES

1. If $A = \begin{bmatrix} 1 & -3 & 2 \\ 4 & 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 4 & 3 & 2 \\ -1 & 0 & 2 \end{bmatrix}$ and $C = \begin{bmatrix} -1 & 4 & -2 \\ 0 & 1 & 0 \end{bmatrix}$,
   determine the elements of the following matrices:
   (a) A + 2B; (b) A + 2B - 3C; (c) $3A^T - B^T$.
2. Remove a common factor from each of the following matrices in order to express it as the product of a number and a matrix:
   (a) $\begin{bmatrix} 8 & -4 \\ -32 & 16 \end{bmatrix}$; (b) $\begin{bmatrix} -x^3 & -x^2 \\ x^2 & -4x^2 \end{bmatrix}$.
3. State the order of the product matrix in each of the following cases:
   (a) $A_{1\times2}.B_{2\times2}$; (b) $A_{3\times1}.B_{1\times2}$; (c) $A_{4\times3}.B_{3\times5}$.
4. For the matrices $A_{2\times2}, B_{2\times3}, C_{3\times3}$ and $D_{2\times4}$, which of the following products are defined.
   (a) A.B; (b) B.C; (c) C.D; (d) A.C; (e) A.D.

5. Determine the elements of the product matrix in each of the following:

(a) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} . \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}$; (b) $\begin{bmatrix} \frac{1}{2} & \frac{1}{3} \\ \frac{1}{4} & 1 \end{bmatrix} . \begin{bmatrix} 12 & 24 \\ 24 & 36 \end{bmatrix}$; (c) $\begin{bmatrix} 0 & 4 & -3 \\ 1 & -1 & 0 \end{bmatrix} . \begin{bmatrix} 2 & 6 \\ 5 & -3 \\ -1 & 7 \end{bmatrix}$;

(d) $\begin{bmatrix} 1 & 2 \\ 3 & 0 \\ 1 & 4 \end{bmatrix} . \begin{bmatrix} 7 & -3 \\ -1 & 4 \end{bmatrix}$; (e) $\begin{bmatrix} 4 & 2 & 1 \end{bmatrix} . \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix}$; (f) $\begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix} . \begin{bmatrix} 4 & 2 & 1 \end{bmatrix}$;

(g) $\begin{bmatrix} -1 & 2 \end{bmatrix} . \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$; (h) $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 6 & 7 \\ 5 & 8 & 9 \end{bmatrix} . \begin{bmatrix} -1 & 3 & -2 \\ -0.5 & -3 & 2.5 \\ 1 & 1 & -1 \end{bmatrix}$.

6. If $\mathrm{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, verify that $\mathrm{A}.\mathrm{A}^{\mathrm{T}} = \mathrm{I}_2$.

7. If $\mathrm{A} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$ and $\mathrm{B} = \begin{bmatrix} -2 & 5 \\ 3 & 4 \end{bmatrix}$, verify that $(\mathrm{A}.\mathrm{B})^{\mathrm{T}} = \mathrm{B}^{\mathrm{T}}.\mathrm{A}^{\mathrm{T}}$ **not** $\mathrm{A}^{\mathrm{T}}.\mathrm{B}^{\mathrm{T}}$.

9.2.6 ANSWERS TO EXERCISES

1. (a) $\begin{bmatrix} 9 & 3 & 6 \\ 2 & 0 & 5 \end{bmatrix}$; (b) $\begin{bmatrix} 12 & -9 & 12 \\ 2 & -3 & 5 \end{bmatrix}$; (c) $\begin{bmatrix} -1 & 13 \\ -12 & 0 \\ 4 & 1 \end{bmatrix}$.

2. (a) $4.\begin{bmatrix} 2 & -1 \\ -8 & 4 \end{bmatrix}$; (b) $-x^2.\begin{bmatrix} x & 1 \\ -1 & 4 \end{bmatrix}$.

3. (a) 1×2; (b) 3×2; (c) 4×5.

4. (a), (b) and (e) are defined.

5. (a) $\begin{bmatrix} 10 & 7 \\ 22 & 15 \end{bmatrix}$; (b) $\begin{bmatrix} 14 & 24 \\ 27 & 42 \end{bmatrix}$; (c) $\begin{bmatrix} 23 & -33 \\ -3 & 9 \end{bmatrix}$;

(d) $\begin{bmatrix} 5 & 5 \\ 21 & -9 \\ 3 & 13 \end{bmatrix}$; (e) $[4]$ defined as 4; (f) $\begin{bmatrix} -4 & -2 & -1 \\ 12 & 6 & 3 \\ 8 & 4 & 2 \end{bmatrix}$;

(g) $\begin{bmatrix} -1 & 2 & -1 \end{bmatrix}$; (h) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

"JUST THE MATHS"

UNIT NUMBER

9.3

MATRICES 3
(Matrix inversion & simultaneous equations)

by

A.J.Hobson

UNIT 9.3 - MATRICES 3

MATRIX INVERSION AND SIMULTANEOUS LINEAR EQUATIONS

9.3.1 INTRODUCTION

In Matrix Algebra, there is no such thing as **division** in the usual sense of this word since we would effectively be dividing by a table of numbers with the headings of the table removed.

However, we may discuss an equivalent operation called **inversion** which is roughly similar to the process in elementary arithmetic where division by a value, a, is the same as multiplication by $\frac{1}{a}$.

For example, consider the solution of a single linear equation in one variable, x, namely

$$mx = k,$$

for which the solution is obviously

$$x = \frac{k}{m}.$$

An apparently over-detailed solution may be set out as follows:

(a) Pre-multiply both sides of the given equation by m^{-1}, giving

$$m^{-1}.(mx) = m^{-1}k.$$

(b) Use the Associative Law of Multiplication in elementary arithmetic to rearrange this as

$$(m^{-1}.m)x = m^{-1}k.$$

(c) Use the property that $m^{-1}.m = 1$ to conclude that

$$1.x = m^{-1}k.$$

(d) Use the fact that the number 1 is the multiplicative identity in elementary arithmetic to conclude that

$$x = m^{-1}k.$$

We shall see later how an almost identical sequence of steps, with matrices instead of numbers, can be used to solve a set of simultaneous linear equations with what is called the **"multiplicative inverse"** of a matrix.

Matrix inversion is a concept which is developed from the rules for matrix multiplication.

9.3.2 MATRIX REPRESENTATION OF SIMULTANEOUS LINEAR EQUATIONS

In this section, we consider the matrix equivalent of three simultaneous linear equations in three unknowns; this case is neither too trivial, nor too difficult to follow. Let the equations have the form:

$$\begin{aligned} a_1x + b_1y + c_1z &= k_1, \\ a_2x + b_2y + c_2z &= k_2, \\ a_3x + b_3y + c_3z &= k_3. \end{aligned}$$

Then, from the properties of matrix equality and matrix multiplication, these can be written as one matrix equation, namely

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} . \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix},$$

which will be written for short in the form

$$\mathrm{MX} = \mathrm{K}.$$

Note:
If we could find a matrix, say N, such that NM = I, we could pre-multiply the above statement by it to give

$$\mathrm{N(MX)} = \mathrm{NK}.$$

That is,

$$\mathrm{(NM)X} = \mathrm{NK}.$$

In other words,

$$\mathrm{IX} = \mathrm{NK}.$$

Hence,

$$\mathrm{X} = \mathrm{NK}.$$

Clearly N exhibits a similar behaviour to the number m^{-1} encountered earlier and, for this reason, a better notation for N would be M^{-1}. This is what we shall use in the work which follows.

9.3.3 THE DEFINITION OF A MULTIPLICATIVE INVERSE

The **"multiplicative inverse"** of a square matrix, M, is another matrix, denoted by M^{-1}, which has the property

$$M^{-1}.M = I.$$

Notes:

(i) It is certainly **possible** for the product of two matrices to be an identity matrix, as illustrated, for example, in the exercises on matrix products in Unit 9.2.

(ii) It is usually acceptable to call M^{-1} the "inverse" of M rather than the "multiplicative inverse", unless we wish to distinguish this kind of inverse from the additive inverse defined in Unit 9.1.

(iii) It can be shown that, when $M^{-1}.M = I$, it is also true that

$$M.M^{-1} = I,$$

even though matrices do not normally commute in multiplication.

(iv) It is easily shown that a square matrix cannot have more than one inverse; for, suppose a matrix A had two inverses, B and C.

Then,

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

which means that C and B are the same matrix.

9.3.4 THE FORMULA FOR A MULTIPLICATIVE INVERSE

So far, we have established the fact that a set of simultaneous linear equations is expressible in the form

$$MX = K$$

and their solution is expressible in the form

$$X = M^{-1}K.$$

We therefore need to establish a method for determining the inverse, M^{-1}, in order to solve the equations for the values of the variables involved.

Development of this method will be dependent on the result known as "Cramer's Rule" for solving simultaneous equations by determinants. (see Units 7.2 and 7.3)

(a) The inverse of a 2 x 2 matrix

Taking

$$M = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}$$

and

$$M^{-1} = \begin{bmatrix} P & Q \\ R & S \end{bmatrix},$$

we require that

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \cdot \begin{bmatrix} P & Q \\ R & S \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Hence,

$$\begin{aligned} a_1 P + b_1 R &= 1, \\ a_2 P + b_2 R &= 0, \\ a_1 Q + b_1 S &= 0, \\ a_2 Q + b_2 S &= 1. \end{aligned}$$

It is easily verified that these equations are satisfied by the solution

$$P = \frac{b_2}{|M|} \quad Q = -\frac{b_1}{|M|} \quad R = -\frac{a_2}{|M|} \quad S = \frac{a_1}{|M|},$$

where

$$|M| = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$

and is called **"the determinant of the matrix** M**"**.

Summary
The formula for the inverse of a 2 x 2 matrix is given by

$$M^{-1} = \frac{1}{|M|} \begin{bmatrix} b_2 & -b_1 \\ -a_2 & a_1 \end{bmatrix},$$

in which we interchange the diagonal elements of M and reverse the sign of the other two elements.

EXAMPLES

1. Write down the inverse of the matrix
$$\mathrm{M} = \begin{bmatrix} 5 & -3 \\ 2 & 7 \end{bmatrix}.$$

   Solution
   Since $\mid \mathrm{M} \mid = 41$, we have
$$\mathrm{M}^{-1} = \frac{1}{41}\begin{bmatrix} 7 & 3 \\ -2 & 5 \end{bmatrix}.$$

   Check
$$\mathrm{M}^{-1}.\mathrm{M} = \frac{1}{41}\begin{bmatrix} 7 & 3 \\ -2 & 5 \end{bmatrix}.\begin{bmatrix} 5 & -3 \\ 2 & 7 \end{bmatrix} = \frac{1}{41}\begin{bmatrix} 41 & 0 \\ 0 & 41 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

2. Use matrices to solve the simultaneous linear equations
$$\begin{aligned} 3x + y &= 1, \\ x - 2y &= 5. \end{aligned}$$

   Solution
   The equations can be written in the form MX = K, where
$$\mathrm{M} = \begin{bmatrix} 3 & 1 \\ 1 & -2 \end{bmatrix}, \quad \mathrm{X} = \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{and} \quad \mathrm{K} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}.$$

   Initially, we **must** check that M is non-singluar by evaluating its determinant:
$$\mid \mathrm{M} \mid = \begin{vmatrix} 3 & 1 \\ 1 & -2 \end{vmatrix} = -\ 6 - 1 = -\ 7.$$

   The inverse matrix is now given by
$$\mathrm{M}^{-1} = -\ \frac{1}{7}\begin{bmatrix} -2 & -1 \\ -1 & 3 \end{bmatrix}.$$

   Thus, the solution of the simultaneous equations is given by
$$\begin{bmatrix} x \\ y \end{bmatrix} = -\ \frac{1}{7}\begin{bmatrix} -2 & -1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 5 \end{bmatrix} = -\ \frac{1}{7}\begin{bmatrix} -7 \\ 14 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

   That is,
$$x = 1 \text{ and } y = -2.$$

Note:
Example 2 should be regarded as a model solution to this type of problem but the student should not expect every exercise to yield whole number answers.

(b) The inverse of a 3 x 3 Matrix

The most convenient version of Cramer's rule to use here may be stated as follows:

The simultaneous linear equations

$$\begin{aligned} a_1x + b_1y + c_1z &= k_1, \\ a_2x + b_2y + c_2z &= k_2, \\ a_3x + b_3y + c_3z &= k_3 \end{aligned}$$

have the solution

$$\frac{x}{\begin{vmatrix} k_1 & b_1 & c_1 \\ k_2 & b_2 & c_2 \\ k_3 & b_3 & c_3 \end{vmatrix}} = \frac{y}{\begin{vmatrix} a_1 & k_1 & c_1 \\ a_2 & k_2 & c_2 \\ a_3 & k_3 & c_3 \end{vmatrix}} = \frac{z}{\begin{vmatrix} a_1 & b_1 & k_1 \\ a_2 & b_2 & k_2 \\ a_3 & b_3 & k_3 \end{vmatrix}} = \frac{1}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}.$$

METHOD
(i) We observe, first, that the last determinant in the Cramer's Rule formula contains the same arrangement of numbers as the matrix M; as in (a), we call it **"the determinant of the matrix** M" and we denote it by | M |.

In this determinant, we let the symbols A_1, A_2, A_3, B_1, B_2, B_3, C_1, C_2 and C_3 denote the **"cofactors"** (or **"signed minors"**) of the elements a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2 and c_3 respectively.

(ii) We now observe that, for each of the elements k_1, k_2 and k_3 in the first three determinants of Cramer's Rule, the cofactor is the same as the cofactor of one of the elements in the final determinant | M |.

(iii) Expanding each of the first three determinants in Cramer's Rule along the column of k-values, the solutions for x, y and z can be written as follows:

$$x = \frac{1}{|\,\mathrm{M}\,|}\left(k_1A_1 + k_2A_2 + k_3A_3\right);$$

$$y = \frac{1}{\mid \mathrm{M} \mid}\left(k_1B_1 + k_2B_2 + k_3B_3\right);$$

$$z = \frac{1}{\mid \mathrm{M} \mid}\left(k_1C_1 + k_2C_2 + k_3C_3\right);$$

or, in matrix format,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{\mid \mathrm{M} \mid}\begin{bmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{bmatrix}\begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix}.$$

Comparing this statement with the statement

$$\mathrm{X} = \mathrm{M}^{-1}K,$$

we conclude that

$$\mathrm{M}^{-1} = \frac{1}{\mid \mathrm{M} \mid}\begin{bmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{bmatrix}.$$

Summary

Since it is known that Cramer's Rule is applicable to any number of equations in the same number of unknowns, similar working would occur for larger or smaller systems of equations. In general, the inverse of a square matrix is **the transpose of the matrix of cofactors times the reciprocal of the determinant of the matrix**.

Notes:

(i) If it should happen that $\mid \mathrm{M} \mid = 0$, then the matrix M does not have an inverse and is said to be **"singular"**. In all other cases, it is said to be **"non-singular"**.

(ii) The transpose of the matrix of cofactors is called the **"adjoint"** of M, denoted by AdjM. There is always an adjoint though not always an inverse; but, when the inverse does exist, we can say that

$$\mathrm{M}^{-1} = \frac{1}{\mid \mathrm{M} \mid}\mathrm{AdjM}.$$

(iii) The inverse of a matrix of order 2×2 fits the above scheme also. The cofactor of each element will be a single number associated with a **"place-sign"** according to the following pattern:

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

Hence, if

$$\mathrm{M} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix},$$

then,

$$\mathrm{M}^{-1} = \frac{1}{a_1 b_2 - a_2 b_1} \begin{bmatrix} b_2 & -b_1 \\ -a_2 & a_1 \end{bmatrix}.$$

As before, we see that the matrix part of the result can be obtained by interchanging the diagonal elements of M and reversing the signs of the other two elements.

EXAMPLE

Use matrices to solve the simultaneous linear equations

$$\begin{aligned} 3x + y - z &= 1, \\ x - 2y + z &= 0, \\ 2x + 2y + z &= 13. \end{aligned}$$

Solution

The equations can be written in the form MX = K, where

$$\mathrm{M} = \begin{bmatrix} 3 & 1 & -1 \\ 1 & -2 & 1 \\ 2 & 2 & 1 \end{bmatrix} \quad \mathrm{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{and} \quad \mathrm{K} = \begin{bmatrix} 1 \\ 0 \\ 13 \end{bmatrix}.$$

Initially, we **must** check that M is non-singluar by evaluating its determinant:

$$\mid \mathrm{M} \mid = \begin{vmatrix} 3 & 1 & -1 \\ 1 & -2 & 1 \\ 2 & 2 & 1 \end{vmatrix} = 3(-2-2) - 1(1-2) + (-1)(2+4) = -17.$$

If C denotes the matrix of cofactors, then

$$\mathrm{C} = \begin{bmatrix} \boxed{-4} & 1 & \boxed{6} \\ -3 & \boxed{5} & -4 \\ \boxed{-1} & -4 & \boxed{-7} \end{bmatrix}.$$

Notes:
(i) The framed elements indicate those for which the place sign is positive and hence no changes of sign are required to convert the minor of the corresponding elements of M into their cofactor. It is a good idea to work these out first.

(ii) The remaining four elements are those for which the place sign is negative and so the cofactor is minus the value of the minor in these cases. If these, too, are worked out together, there is less scope for making mistakes.

(iii) It is important to remember that, in finding the elements of C, we **do not multiply the cofactors of the elements in** M **by the elements themselves**.

The transpose of C now provides the adjoint matrix and hence the inverse as follows:

$$\mathrm{M}^{-1} = \frac{1}{\mid \mathrm{M} \mid}\mathrm{AdjM} = \frac{1}{-17}\mathrm{C}^{\mathrm{T}} = \frac{1}{-17}\begin{bmatrix} -4 & -3 & -1 \\ 1 & 5 & -4 \\ 6 & -4 & -7 \end{bmatrix}.$$

Thus, the solution of the simultaneous equations is given by

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{-17}\begin{bmatrix} -4 & -3 & -1 \\ 1 & 5 & -4 \\ 6 & -4 & -7 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 13 \end{bmatrix} = \frac{1}{-17}\begin{bmatrix} -17 \\ -51 \\ -85 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}.$$

That is,

$$x = 1 \quad y = 3 \quad \text{and} \quad z = 5.$$

Note:
Once again, the example should be regarded as a model solution to this type of problem but the student should not expect every exercise to yield whole number answers.

9.3.5 EXERCISES

1. Determine, where possible, the multiplicative inverses of the following matrices:

(a) $\begin{bmatrix} 1 & 5 \\ 7 & 9 \end{bmatrix}$; (b) $\begin{bmatrix} -2 & 1 \\ -6 & 3 \end{bmatrix}$; (c) $\begin{bmatrix} -1 & 2 \\ 1 & 2 \end{bmatrix}$;

(d) $\begin{bmatrix} 1 & 4 & -2 \\ -3 & 2 & -7 \\ 1 & 0 & 5 \end{bmatrix}$; (e) $\begin{bmatrix} 6 & 3 & 1 \\ 0 & -5 & 2 \\ 1 & 1 & 4 \end{bmatrix}$; (f) $\begin{bmatrix} 11 & 3 & 5 \\ 6 & 2 & 2 \\ 10 & 3 & 4 \end{bmatrix}$.

2. Use a matrix inverse method to solve the following sets of simultaneous linear equations:

(a)

$$\begin{aligned} 4x - 3y &= 9, \\ 3x - 2y &= 7; \end{aligned}$$

(b)

$$\begin{aligned} x + 2y &= -2, \\ 5x - 4y &= 3; \end{aligned}$$

(c)

$$\begin{aligned} 3x + 2y + z &= 3, \\ 5x + 4y + 3z &= 3, \\ 6x + y + z &= 5; \end{aligned}$$

(d)

$$\begin{aligned} x + 2y + 2z &= 0, \\ 2x + 5y + 4z &= 2, \\ x - y - 6z &= 4; \end{aligned}$$

(e)

$$\begin{aligned} 3x + 2y - 2z &= 16, \\ 4x + 3y + 3z &= 2' \\ 2x - y + z &= -1; \end{aligned}$$

(f)

$$\begin{aligned} 2(i_3 - i_2) + 5(i_3 - i_1) &= 4, \\ (i_2 - i_3) + 2i_2 + (i_2 - i_1) &= 0, \\ 5(i_1 - i_3) + 2(i_1 - i_2) + i_2 &= 1. \end{aligned}$$

3. If A $= \begin{bmatrix} 4 & 0 & 1 \\ 2 & 1 & 3 \\ -1 & -2 & 5 \end{bmatrix}$ and C $= \begin{bmatrix} -1 & 2 & 1 \\ 3 & 1 & -4 \\ 0 & 5 & -2 \end{bmatrix}$,
determine a matrix, B, such that A.B = C.

9.3.6 ANSWERS TO EXERCISES

1. (a) $-\frac{1}{26}\begin{bmatrix} 9 & -5 \\ -7 & 1 \end{bmatrix}$; (b) No Inverse, (c) $-\frac{1}{4}\begin{bmatrix} 2 & -2 \\ -1 & -1 \end{bmatrix}$;
   (d) $\frac{1}{46}\begin{bmatrix} 10 & -20 & -24 \\ 8 & 7 & 13 \\ -2 & 4 & 14 \end{bmatrix}$; (e) $-\frac{1}{121}\begin{bmatrix} -22 & -11 & 11 \\ 2 & 23 & -12 \\ 5 & -3 & -30 \end{bmatrix}$; (f) There is no inverse.
2. (a) $x = 3$ and $y = 1$; (b) $x = \frac{1}{7}$ and $y = -\frac{13}{14}$; (c) $x = 1$, $y = 1$ and $z = -2$;
   (d) $x = -\frac{3}{2}$, $y = 2$ and $z = -\frac{5}{4}$; (e) $x = 2$, $y = \frac{3}{2}$ and $z = -\frac{7}{2}$;
   (f) $i_1 = \frac{11}{6}$, $i_2 = 1$ and $i_3 = \frac{13}{6}$.
3.

$$B = \frac{1}{41}\begin{bmatrix} 11 & -2 & -1 \\ -13 & 21 & -10 \\ -3 & 8 & 4 \end{bmatrix} . \begin{bmatrix} -1 & 2 & 1 \\ 3 & 1 & -4 \\ 0 & 5 & -2 \end{bmatrix} = \frac{1}{41}\begin{bmatrix} -17 & 15 & 21 \\ 76 & -55 & -77 \\ 27 & 22 & -43 \end{bmatrix}.$$

"JUST THE MATHS"

UNIT NUMBER

9.4

MATRICES 4
(Row operations)

by

A.J.Hobson

UNIT 9.4 - MATRICES 4 - ROW OPERATIONS

9.4.1 MATRIX INVERSES BY ROW OPERATIONS

In this section, we shall examine an alternative method for finding the multiplicative inverse of a matrix but the techniques introduced will lead on to other aspects of solving simultaneous linear equations not discussed in earlier units.

DEFINITION
An **"elementary row operation"** on a matrix is any one of the following three possibilities:

(a) The interchange of two rows;
(b) The multiplication of the elements in any row by a non-zero number;
(c) The addition of the elements in any row to the corresponding elements in another row.

Notes:
(i) Elementary row operations are essentially the same kind of operations as those used in the solution of a set of simultaneous linear equations by the method of elimination; but here we shall be considering sets of coefficients in the form of matrices rather than complete sets of equations.

(ii) Elementary row operations of types (b) and (c) imply that the elements in any row may be **subtracted** from the corresponding elements in another row and, more generally, multiples of the elements in any row may be added to or subtracted from the corresponding elements in another row.

RESULT 1.
To perform an elementary row operation on a matrix **algebraically**, we may pre-multiply the matrix by an identity matrix on which the same elementary row operation has been already performed. For example, in the matrix

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix},$$

suppose we wished to subtract twice the third row from the second row. It is easy enough to carry this out by inspection, but could also be regarded as the succession of two elementary row operations as follows:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} . \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} . \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} . \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ 2a_3 & 2b_3 \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ a_2 - 2a_3 & b_2 - 2b_3 \\ a_3 & b_3 \end{bmatrix}.$$

DEFINITION
An **"elementary matrix"** is a matrix obtained from an identity matrix by performing upon it one elementary row operation.

RESULT 2.
If a certain sequence of elementary row operations converts a given square matrix, M, into the corresponding identity matrix, then the same sequence of elementary row operations in the same order will convert the identity matrix into M^{-1}.

Proof:

Suppose that

$$E_n.E_{n-1}....E_4.E_3.E_2.E_1.M = \mathrm{I}$$

where $E_1, E_2, E_3, E_4, ..., E_{n-1}, En$ are elementary matrices.

Then, by post-multiplying both sides with M^{-1}, we obtain

$$E_n.E_{n-1}....E_4.E_3.E_2.E_1.\mathrm{M}.\mathrm{M}^{-1} = \mathrm{I}.\mathrm{M}^{-1}.$$

In other words,

$$E_n.E_{n-1}....E_4.E_3.E_2.E_1.\mathrm{I} = \mathrm{M}^{-1},$$

which proves the result.

EXAMPLES

1. Use elementary row operations to determine the inverse of the matrix

$$\begin{bmatrix} 3 & 7 \\ 2 & 5 \end{bmatrix}.$$

Solution

First we write down the given matrix side-by-side with the corresponding identity matrix in the following format:

$$\begin{bmatrix} 3 & 7 & \vdots & 1 & 0 \\ 2 & 5 & \vdots & 0 & 1 \end{bmatrix}.$$

Secondly, we try to arrange that the first element in the first column of this arrangement is 1; and this can be carried out by subtracting the second row from the first row. (Instruction: $R_1 \to R_1 - R_2$).

$$\left[\begin{array}{cccc} 1 & 2 & \vdots & 1 & -1 \\ 2 & 5 & \vdots & 0 & 1 \end{array}\right].$$

Thirdly, we try to convert the first column of the display into the first column of the identity matrix; and this can be carried out by subtracting twice the first row from the second row.
(Instruction: $R_2 \to R_2 - 2R_1$).

$$\left[\begin{array}{ccccc} 1 & 2 & \vdots & 1 & -1 \\ 0 & 1 & \vdots & -2 & 3 \end{array}\right].$$

Lastly, we try to convert the second column of the display into the second column of the identity matrix: and this can be carried out by subtracting twice the second row from the first row.
(Instruction: $R_1 \to R_1 - 2R_2$).

$$\left[\begin{array}{ccccc} 1 & 0 & \vdots & 5 & -7 \\ 0 & 1 & \vdots & -2 & 3 \end{array}\right].$$

The inverse matrix is therefore

$$\left[\begin{array}{cc} 5 & -7 \\ -2 & 3 \end{array}\right],$$

as we would have obtained by the cofactor method.

Notes:

(i) The technique used for a 2×2 matrix applies to square matrices of all orders with appropriate modifications.

(ii) The idea of the method is to obtain, in chronological order, the columns of the identity matrix from the columns of the given matrix. We do this by using elementary row operations to convert each diagonal element in the given matrix to 1 and then using multiples of 1 to reduce the remaining elements in the same column to zero.

(iii) Once any row has been used to reduce elements to zero, that row must not be used again as the operator; otherwise the zeros obtained may change to other values.

2. Use elementary row operations to show that the matrix

$$\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$$

has no inverse.

Solution

We set up the scheme in the following format:

$$\begin{bmatrix} 1 & 3 & \vdots & 1 & 0 \\ 2 & 6 & \vdots & 0 & 1 \end{bmatrix}.$$

Then, we proceed according to the instructions indicated:

$R_2 \to R_2 - 2R_1$

$$\begin{bmatrix} 1 & 3 & \vdots & 1 & 0 \\ 0 & 0 & \vdots & -1 & 1 \end{bmatrix}.$$

There is no way now of continuing to convert the given matrix into the identity matrix; hence, there is no inverse.

3. Use elementary row operations to determine the inverse of the matrix

$$\begin{bmatrix} 4 & 1 & 6 \\ 2 & 1 & 3 \\ 3 & 2 & 5 \end{bmatrix}.$$

Solution

We set up the scheme in the following format:

$$\begin{bmatrix} 4 & 1 & 6 & \vdots & 1 & 0 & 0 \\ 2 & 1 & 3 & \vdots & 0 & 1 & 0 \\ 3 & 2 & 5 & \vdots & 0 & 0 & 1 \end{bmatrix}.$$

Then, we proceed according to the instructions indicated:

$R_1 \to R_1 - R_3$

$$\begin{bmatrix} 1 & -1 & 1 & \vdots & 1 & 0 & 1 \\ 2 & 1 & 3 & \vdots & 0 & 1 & 0 \\ 3 & 2 & 5 & \vdots & 0 & 0 & 1 \end{bmatrix};$$

$R_2 \to R_2 - 2R_1$ and $R_3 \to R_3 - 3R_1$

$$\begin{bmatrix} 1 & -1 & 1 & \vdots & 1 & 0 & -1 \\ 0 & 3 & 1 & \vdots & -2 & 1 & 2 \\ 0 & 5 & 2 & \vdots & -3 & 0 & 4 \end{bmatrix};$$

$R_2 \to 2R_2 - R_3$

$$\begin{bmatrix} 1 & -1 & 1 & \vdots & 1 & 0 & -1 \\ 0 & 1 & 0 & \vdots & -1 & 2 & 0 \\ 0 & 5 & 2 & \vdots & -3 & 0 & 4 \end{bmatrix};$$

$R_1 \to R_1 + R_2$ and $R_3 \to R_3 - 5R_2$

$$\begin{bmatrix} 1 & 0 & 1 & \vdots & 0 & 2 & -1 \\ 0 & 1 & 0 & \vdots & -1 & 2 & 0 \\ 0 & 0 & 2 & \vdots & 2 & -10 & 4 \end{bmatrix};$$

$R_3 \to R_3 \times \frac{1}{2}$

$$\begin{bmatrix} 1 & 0 & 1 & \vdots & 0 & 2 & -1 \\ 0 & 1 & 0 & \vdots & -1 & 2 & 0 \\ 0 & 0 & 1 & \vdots & 1 & -5 & 2 \end{bmatrix};$$

$R_1 \to R_1 - R_3$

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

The required inverse matrix is therefore

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

9.4.2 GAUSSIAN ELIMINATION - THE ELEMENTARY VERSION

Elementary row operations can also be conveniently used in another method of solving simultaneous linear equations which relates closely again to the elimination method sometimes encountered in courses which do not include matrices. The method will be introduced through the case of three equations in three unkowns, but may be applied in other cases as well.

Suppose a set of simultaneous linear equations in the variables x, y and z appeared in the special form:

$$
\begin{aligned}
a_1x + b_1y + c_1z &= k_1, \\
b_2y + c_2z &= k_2, \\
c_3z &= k_3.
\end{aligned}
$$

Then it is very simple to solve the equations by first obtaining z from the third equation, substituting its value into the second equation in order to find y, then substituting for both y and z in the first equation in order to find x.

The method of Gaussian Elimination reduces any set of linear equations to this triangular form by adding or subtracting suitable multiples of pairs of the equations; but the method is more conveniently laid out in a tabular form using only the coefficients of the variables and the constant terms. We illustrate with an example.

EXAMPLE
Solve the simultaneous linear equations

$$
\begin{aligned}
2x + y + z &= 3, \\
x - 2y - z &= 2, \\
3x - y + z &= 8.
\end{aligned}
$$

Solution
For the simplest arithmetic, we try to arrange that the first coefficient in the first equation is 1. In the current example, we could interchange the first two equations.

$$
\begin{array}{ccc|c}
\boxed{1} & -2 & -1 & 2 \\
2 & 1 & 1 & 3 \\
3 & -1 & 1 & 8
\end{array}
$$

This format is known as an **"augmented matrix"**. The matrix of coefficients has been augmented by the matrix of constant terms.

Using the notation of the previous section, we apply the instructions $R_2 \rightarrow R_2 - 2R_1$ and $R_3 \rightarrow R_3 - 3R_1$ giving a new table, namely

$$\begin{array}{rrr|r} 1 & -2 & -1 & 2 \\ 0 & \boxed{5} & 3 & -1 \\ 0 & 5 & 4 & 2 \end{array}$$

Next, we apply the instruction $R_3 \rightarrow R_3 - R_2$ giving

$$\begin{array}{rrr|r} 1 & -2 & -1 & 2 \\ 0 & 5 & 3 & -1 \\ 0 & 0 & 1 & 3 \end{array}$$

The numbers enclosed in the boxes are called the **"pivot elements"** and are used to reduce to zero the elements below them in the same column.

The final table above provides a new set of equations, equivalent to the original, namely

$$\begin{aligned} x - 2y - z &= 2, \\ 5y + 3z &= -1, \\ z &= 3. \end{aligned}$$

Hence, $\boxed{z = 3, \quad y = -2, \quad x = 1}$.

INSERTING A CHECK COLUMN

In the above example, the numbers were fairly simple, giving little scope for careless mistakes. However, with large numbers of equations, often involving awkward decimal quantities, the margin for error is greatly increased.

As a check on the arithmetic at each stage, we deliberately introduce some **additional** arithmetic which has to remain consistent with the calculations already being carried out. The method is to add together the numbers in each row in order to produce an extra column; each row operation is then performed on the extended rows with the result that, in the new table, the final column should still be the sum of the numbers to the left of it.

The working for our previous example would be as follows:

$$\begin{array}{ccc|c|c} \boxed{1} & -2 & -1 & 2 & 0 \\ 2 & 1 & 1 & 3 & 7 \\ 3 & -1 & 1 & 8 & 11 \end{array}$$

$R_2 \to R_2 - 2R_1, R_3 \to R_3 - 3R_1$

$$\begin{array}{ccc|c|c} 1 & -2 & -1 & 2 & 0 \\ 0 & \boxed{5} & 3 & -1 & 7 \\ 0 & 5 & 4 & 2 & 11 \end{array}$$

$R_3 \to R_3 - R_2$

$$\begin{array}{ccc|c|c} 1 & -2 & -1 & 2 & 0 \\ 0 & 5 & 3 & -1 & 7 \\ 0 & 0 & 1 & 3 & 4 \end{array}$$

9.4.4 EXERCISES

1. Use elementary row operations to determine (where possible) the inverses of the following matrices:

   (a) $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$;

   (b) $\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$;

   (c) $\begin{bmatrix} 2 & 4 \\ 3 & 5 \end{bmatrix}$;

   (d) $\begin{bmatrix} 1 & 2 & -1 \\ 2 & 0 & -7 \\ 3 & 11 & 3 \end{bmatrix}$;

   (e) $\begin{bmatrix} 0 & 3 & -2 \\ 1 & -1 & 5 \\ 1 & 5 & 1 \end{bmatrix}$;

   (f) $\begin{bmatrix} 2 & -3 & 4 \\ -1 & 3 & -4 \\ 1 & 0 & 2 \end{bmatrix}$;

   (g) $\begin{bmatrix} -8 & -7 & -6 & -5 \\ -4 & -3 & -2 & -1 \\ 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}$.

2. Use Gaussian Elimination (with a check column) to solve the following sets of simultaneous linear equations:

   (a)
$$\begin{aligned} x + 3y &= -8, \\ 5x - 2y &= 11; \end{aligned}$$

   (b)
$$\begin{aligned} x + 2y &= -2, \\ 5x - 4y &= 3; \end{aligned}$$

   (c)
$$\begin{aligned} 2x - y + z &= 7, \\ 3x + y - 5z &= 13, \\ x + y + z &= 5; \end{aligned}$$

   (d)
$$\begin{aligned} 3x + 2y - 2z &= 16, \\ 4x + 3y + 3z &= 2, \\ 2x - y + z &= -1. \end{aligned}$$

9.4.5 ANSWERS TO EXERCISES

1. (a) $\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$;

   (b) $\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$;

   (c) $\begin{bmatrix} -\frac{5}{2} & 2 \\ \frac{3}{2} & -1 \end{bmatrix}$;

   (d) $\begin{bmatrix} 77 & -17 & -14 \\ -27 & 6 & 5 \\ 22 & -5 & -4 \end{bmatrix}$;

   (e) There is no inverse;

   (f) $\begin{bmatrix} 1 & 1 & 0 \\ -\frac{1}{3} & 0 & \frac{2}{3} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$;

   (g) There is no inverse.

2. (a) $x = 1$ and $y = -3$;

   (b) $x = -\frac{1}{7}$ and $y = -\frac{13}{14}$;

   (c) $x = 4$, $y = 1$ and $z = 0$;

   (d) $x = 2$, $y = \frac{3}{2}$ and $z = -\frac{7}{2}$.

"JUST THE MATHS"

UNIT NUMBER

9.5

MATRICES 5
(Consistency and rank)

by

A.J.Hobson

UNIT 9.5 - MATRICES 5 - CONSISTENCY AND RANK

9.5.1 THE CONSISTENCY OF SIMULTANEOUS LINEAR EQUATIONS

Introduction

Methods of solving simultaneous linear equations in earlier Units have already shown that some sets of equations cannot be solved to give a unique solution. The Gaussian Elimination method described in Unit 9.4 is able to detect such situations as illustrated in the work below:

ILLUSTRATION 1.
Suppose, firstly, that we were required to investigate the solution of the simultaneous linear equations

$$\begin{aligned} 3x - y + z &= 1, \\ 2x + 2y - 5z &= 0, \\ 5x + y - 4z &= 7. \end{aligned}$$

The Gaussian Elimination solution, with check column, proceeds as follows

$$\begin{array}{ccc|c|c} 3 & -1 & 1 & 1 & 4 \\ 2 & 2 & -5 & 0 & -1 \\ 5 & 1 & -4 & 7 & 9 \end{array}$$

$R_1 \rightarrow R_1 - R_2$

$$\begin{array}{ccc|c|c} 1 & -3 & 6 & 1 & 5 \\ 2 & 2 & -5 & 0 & -1 \\ 5 & 1 & -4 & 7 & 9 \end{array}$$

$R_2 \rightarrow R_2 - 2R_1$ and $R_3 \rightarrow R_3 - 5R_1$

$$\begin{array}{ccc|c|c} 1 & -3 & 6 & 1 & 5 \\ 0 & 8 & -17 & -2 & -11 \\ 0 & 16 & -34 & 2 & -16 \end{array}$$

$R_3 \rightarrow R_3 - 2R_2$

$$\begin{array}{ccc|c|c} 1 & -3 & 6 & 1 & 5 \\ 0 & 8 & -17 & -2 & -11 \\ 0 & 0 & 0 & 6 & 6 \end{array}$$

The third line seems to imply that $0.z = 6$; that is, $0 = 6$ which is impossible.

Hence the equations have no solution and are said to be **"inconsistent"**.

ILLUSTRATION 2.
Secondly, let us consider the simultaneous linear equations

$$\begin{aligned} 3x - y + z &= 1, \\ 2x + 2y - 5z &= 0, \\ 5x + y - 4z &= 1, \end{aligned}$$

which differ from the first illustration in one number only, the constant term of the third equation; but the conclusion will be very different.

This time, the Gaussian Elimination method gives the following sequence of steps:

$$\begin{array}{lll|l|l} 3 & -1 & 1 & 1 & 4 \\ 2 & 2 & -5 & 0 & -1 \\ 5 & 1 & -4 & 1 & 3 \end{array}$$

$R_1 \rightarrow R_1 - R_2$

$$\begin{array}{lll|l|l} 1 & -3 & 6 & 1 & 5 \\ 2 & 2 & -5 & 0 & -1 \\ 5 & 1 & -4 & 1 & 3 \end{array}$$

$R_2 \rightarrow R_2 - 2R_1$ and $R_3 \rightarrow R_3 - 5R_1$

$$\begin{array}{lll|l|l} 1 & -3 & 6 & 1 & 5 \\ 0 & 8 & -17 & -2 & -11 \\ 0 & 16 & -34 & -4 & -22 \end{array}$$

$R_3 \rightarrow R_3 - 2R_2$

$$\begin{array}{lll|l|l} 1 & -3 & 6 & 1 & 5 \\ 0 & 8 & -17 & -2 & -11 \\ 0 & 0 & 0 & 0 & 0 \end{array}$$

The third line here implies that the third equation is redundant since any set of x, y and z values would satisfy it. Hence, the original set of equations is equivalent to two equations in three unknowns, namely

$$\begin{aligned} x - 3y + 6z &= 1, \\ 8y - 17z &= -2. \end{aligned}$$

To state a suitable form of the conclusion, we could use the fact that any one of the three variables may be chosen at random, the other two being expressed in terms of it. For instance, if we choose z at random, then

$$x = \frac{3z+2}{8} \quad \text{and} \quad y = \frac{17z-2}{8}.$$

However, there is a neater way of arriving at the conclusion.

Neater form of solution

Suppose that $x = x_0$, $y = y_0$, $z = z_0$ is any **known** solution to the equations. It could be determined, for instance, by starting with $z = 0$; in this case, $z = 0$, $y = -\frac{1}{4}$ and $x = \frac{1}{4}$.

Let us now substitute

$$\begin{aligned} x &= x_1 + x_0, \\ y &= y_1 + y_0, \\ z &= z_1 + z_0, \end{aligned}$$

from which we obtain

$$\begin{aligned} (x_1 + x_0) - 3(y_1 + y_0) + 6(z_1 + z_0) &= 1, \\ 8(y_1 + y_0) - 17(z_1 + z_0) &= -2. \end{aligned}$$

But, because (x_0, y_0, z_0) is a known solution, this reduces to

$$\begin{aligned} x_1 - 3y_1 + 6z_1 &= 0, \\ 8y_1 - 17z_1 &= 0. \end{aligned}$$

This is a set of **"homogeneous linear equations"** and, although clearly satisfied by $x_1 = 0$, $y_1 = 0$, $z_1 = 0$, we regard this as a **"trivial solution"** and ignore it.

Of more use to us is the fact that an infinite number of non-trivial solutions can be found for each of which the variables x_1, y_1 and z_1 are in a certain set of ratios. In the present case, from the second equation, we have

$y_1 = \frac{17}{8}z_1$ which means that $y_1 : z_1 = 17 : 8$.

Substituting into the first equation gives

$x_1 - \frac{51}{8}z_1 + 6z_1 = 0$ which means that $x_1 = \frac{3}{8}z_1$; that is, $x_1 : z_1 = 3 : 8$.

Combining these conclusions, we can say that

$$x_1 : y_1 : z_1 = 3 : 17 : 8$$

and any three numbers in these ratios will serve as a set of values for x_1, y_1 and z_1.

The neater form of solution can, in general, be written

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \alpha \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix},$$

where α may be any non-zero number.

In our present example,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{4} \\ -\frac{1}{4} \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} 3 \\ 17 \\ 8 \end{bmatrix}.$$

Notes:
(i) An alternative way of determining the ratios $x_1 : y_1 : z_1$ is to use the fact that, since at least one of the variables is going to be non-zero, we may begin by letting it equal 1.

In the above illustration, suppose we let $z_1 = 1$; then

$$y_1 = \frac{17}{8} \quad \text{and} \quad x_1 = \frac{3}{8}.$$

Hence,

$$x_1 : y_1 : z_1 = \frac{3}{8} : \frac{17}{8} : 1,$$

which can be rewritten

$$x_1 : y_1 : z_1 = 3 : 17 : 8,$$

as before.

(ii) Should it happen that a set of simultaneous linear equations reduces to **only one** equation (that is, each equation is just a multiple of the first) then a similar procedure can be applied as in the following illustration:

ILLUSTRATION 3.

Using trial and error, the equation

$$3x - 2y + 5z = 6$$

has a particular solution $x_0 = 1, y_0 = 1, z_0 = 1$, so that the general solution is given by

$$x = x_0 + x_1,\ \ y = y_0 + y_1,\ \ z = z_0 + z_1,$$

where

$$3x_1 - 2y_1 + 5z_1 = 0.$$

We could choose $x_1 = \alpha$ and $y_1 = \beta$ at random for this equation to give $z_1 = \frac{2\beta - 3\alpha}{5}$. That is,

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 0 \\ -\frac{3}{5} \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ \frac{2}{5} \end{bmatrix}.$$

Hence,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \alpha \begin{bmatrix} 1 \\ 0 \\ -\frac{3}{5} \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ \frac{2}{5} \end{bmatrix}.$$

9.5.2 THE ROW-ECHELON FORM OF A MATRIX

In the Gaussian Elimination method to solve a set of simltaneous linear equations,

$$MX = K,$$

we begin with the augmented matrix M|K and use elementary row operations to obtain more zeros at the beginning of each row than at the beginning of the previous row.

If desired, the first non-zero element in each row could be reduced to 1 by simply dividing that row throughout by the value of the first non-zero element. This leads to the following definition:

DEFINITION
The **"row echelon form"** of a matrix is that for which the first non-zero element in each row is 1 and occurs to the right of the first non-zero element in the previous row.

Note:
When using the term "row echelon form" in future, we shall not insist that the first non-zero element in each row has **actually** been reduced to 1.

9.5.3 THE RANK OF A MATRIX

The two illustrations of Gaussian Elimination discussed in section 9.5.1 may be used to imply that special conclusions are reached when, in the final row echelon form, either a complete row of M has reduced to zero (Illustration 1.) or a complete row of M|K has reduced to zero (Illustration 2.) This leads to another definition as follows:

DEFINITION
The **"rank"** of a matrix is the number or rows which do not reduce to a complete row of zeros when the matrix has been converted to row echelon form.

ILLUSTRATIONS

1. In our previous Illustration 1, M had rank 2 but M|K had rank 3. The equations were inconsistent.
2. In our previous Illustration 2, M had rank 2 and M|K also had rank 2. The equations had an infinite number of solutions.
3. In the examples of Unit 9.4 both M and M|K had rank 3. There was a unique solution to the simultaneous equations.

A general summary of these observations may be set out as follows:

1. The equations MX = K are inconsistent if rank M $<$ rank M|K.
2. The equations MX = K have an infinite number of solutions if rank M = rank M|K $< n$ where n is the number of equations.
3. The equations MX = K have a unique solution if rank M = rank M|K $= n$ where n is the number of equations.

9.5.4 EXERCISES

1. Use Gaussian Elimination to show that the following sets of simultaneous equations are inconsistent:

   (a)

$$\begin{aligned} x - y + 2z &= 2, \\ 3x + y - z &= 3, \\ 5x - y + 3z &= 4; \end{aligned}$$

   (b)

$$\begin{aligned} x - y + 2z &= 1, \\ -x + 3y - z &= -1, \\ 3x - 7y + 4z &= 5. \end{aligned}$$

2. Determine the rank of the following matrices:

   (a) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 5 \\ 3 & 4 & 11 & 2 \end{bmatrix}$; (b) $\begin{bmatrix} a & 1 & 2 \\ -1 & -3 & 8 \\ 1 & 12 & -3 \\ 4 & -3 & 7 \end{bmatrix}$.

3. Determine the general solution of the following equations by reducing the augmented matrix to row echelon form:

$$\begin{aligned} x + 3y - z &= 6, \\ 8x + 9y + 4z &= 21, \\ 2x + y + 2z &= 3. \end{aligned}$$

4. State the value of t for which the matrix

$$\mathrm{M} = \begin{bmatrix} 2 & 1 & -3 \\ 4 & t & -6 \\ 3t & 3 & -9 \end{bmatrix}$$

has rank 1 and determine the general solution of the system of equations

$$\mathrm{MX} = \mathrm{K},$$

where

$$\mathrm{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \text{ and } \mathrm{K} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

5. For the system of simultaneous linear equations

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 1, \\ 2x_1 - x_2 + tx_3 &= 2, \\ -x_1 + 2x_2 + x_3 &= s, \end{aligned}$$

determine for which values of s and t there exists

(a) no solution;

(b) a unique solution;

(c) an infinite number of solutions.

Solve the equations for the two cases $s = 1, t = 1$ and $s = -1, t = 7$.

6. Determine the values of t for which the matrix

$$\mathrm{M} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & t \\ 3t & 2 & -2 \end{bmatrix}$$

has rank 2.

For each of these two values of t, solve the system of equations

$$\mathrm{MX} = \mathrm{K},$$

where $\mathrm{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and $\mathrm{K} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$.

9.5.5 ANSWERS TO EXERCISES

1. The rank of the matrix of x, y and z coefficients is less than the rank of the augmented matrix.

2. (a) 2; (b) 3.

3. $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{3}{5} \\ \frac{9}{5} \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -7 \\ 4 \\ 5 \end{bmatrix}.$

4.
$$t = 2$$

and

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \alpha \begin{bmatrix} 1 \\ 0 \\ \frac{2}{3} \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ \frac{1}{3} \end{bmatrix}.$$

5. (a) $t = 7$ and $s \neq -1$; (b) $t \neq 7$; (c) $t = 7$ and $s = -1$.
If $s = 1$ and $t = -1$, then $x_1 = \frac{7}{4}$, $x_2 = \frac{5}{4}$ and $x_3 = \frac{1}{4}$.
If $s = -1$ and $t = 7$, then
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} 5 \\ 3 \\ -1 \end{bmatrix}.$$

6. The rank is 2 when $t = -1$ or when $t = \frac{2}{3}$.
If $t = -1$, the equations are inconsistent.
If $t = \frac{2}{3}$, the equations have general solution
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -5 \\ 8 \\ 3 \end{bmatrix}.$$

"JUST THE MATHS"

UNIT NUMBER

9.6

MATRICES 6
(Eigenvalues and eigenvectors)

by

A.J.Hobson

UNIT 9.6 - MATRICES 6

EIGENVALUES AND EIGENVECTORS

9.6.1 THE STATEMENT OF THE PROBLEM

Suppose A is any square matrix, and let X be a column vector with the same number of rows as there are columns in A. For example, if A is of order $m \times m$, then X must be of order $m \times 1$ and AX will also be of order $m \times 1$.

We ask the question:

"Is it ever possible that AX can be just a scalar multiple of X ?"

We exclude the case when the elements of X are all zero since, in a practical application, these elements will usally be the components of an actual vector quantity; and, if they are all zero, the direction of the vector will be indeterminate.

ILLUSTRATIONS

1.

$$\begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 7 \\ -2 \\ -3 \end{bmatrix} = \begin{bmatrix} 7 \\ -2 \\ -3 \end{bmatrix}.$$

2.

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

The formal statement of the problem

For a given square matrix, A, we investigate the existence of column vectors, X, such that

$$\mathrm{AX} = \lambda \mathrm{X}$$

for some scalar quantity λ. Each such column vector is called an **"eigenvector"** of the matrix, A; and each corresponding value of λ is called an **"eigenvalue"** of the matrix, A.

Notes:

(i) The German word "eigen" (meaning "hidden") gives rise to the above names, but other alternatives are "latent values and latent vectors" or "characteristic values and characteristic vectors".

(ii) In the discussion which follows, A will be, mostly, a matrix of order 3×3, but the ideas involved will apply to square matrices of other orders also (see Example 1 and Exercises 9.6.3, question 1).

9.6.2 THE SOLUTION OF THE PROBLEM

Assuming that

$$\mathrm{A} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \text{ and } \mathrm{X} = \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

the matrix equation, $\mathrm{AX} = \lambda\mathrm{X}$, may be written out fully in the following form:

$$\begin{aligned} a_1x + b_1y + c_1z &= \lambda x, \\ a_2x + b_2y + c_2z &= \lambda y, \\ a_3x + b_3y + c_3z &= \lambda z; \end{aligned}$$

or, on rearrangement,

$$\begin{aligned} (a_1 - \lambda)x + b_1y + c_1z &= 0, \\ a_2x + (b_2 - \lambda)y + c_2z &= 0, \\ a_3x + b_3y + (c_3 - \lambda)z &= 0, \end{aligned}$$

which is a set of homogeneous linear equations in x, y and z and may be written, for short, in the form

$$(\mathrm{A} - \lambda\mathrm{I})\mathrm{X} = [0],$$

where I denotes the identity matrix of order 3×3.

From the results of Unit 7.4, the condition that the three homogeneous linear equations have a solution other than $x = 0$, $y = 0$, $z = 0$ is

$$|A - \lambda I| = \begin{vmatrix} a_1 - \lambda & b_1 & c_1 \\ a_2 & b_2 - \lambda & c_2 \\ a_3 & b_3 & c_3 - \lambda \end{vmatrix} = 0.$$

This equation will give, on expansion, a cubic equation in λ called the **"characteristic equation"** of A. Its left-hand side is called the **"characteristic polynomial"** of A.

The characteristic equation of a 3×3 matrix, being a cubic equation, will (in general) have three solutions.

EXAMPLES

1. Determine the eigenvalues and eigenvectors of the matrix

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

Solution

(a) The eigenvalues

The characteristic equation is given by

$$0 = |A - \lambda I| = \begin{vmatrix} 2 - \lambda & 4 \\ 5 & 3 - \lambda \end{vmatrix} = \lambda^2 - 5\lambda - 14 = (\lambda + 2)(\lambda - 7).$$

The eigenvalues are therefore $\lambda = -2$ and $\lambda = 7$.

(b) The eigenvectors

Case 1. $\lambda = -2$

We require to solve the equation $x + y = 0$,
giving $x : y = -1 : 1$ and a corresponding eigenvector

$$X = \alpha \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

where α is any **non-zero** scalar.

Case 2. $\lambda = 7$

We require to solve the equation $5x - 4y = 0$,
giving $x : y = 4 : 5$ and a corresponding eigenvector

$$X = \beta \begin{bmatrix} 4 \\ 5 \end{bmatrix},$$

where β is any **non-zero** scalar.

2. Determine the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}.$$

Solution

(a) The eigenvalues

The characteristic equation is given by

$$0 = |A - \lambda I| = \begin{vmatrix} 3-\lambda & 2 & 4 \\ 2 & -\lambda & 2 \\ 4 & 2 & 3-\lambda \end{vmatrix}.$$

Direct expansion of the determinant gives the equation

$$-\lambda^3 + 6\lambda^2 + 15\lambda + 8 = 0,$$

which will factorise into

$$(1+\lambda)^2(8-\lambda) = 0.$$

Note:

Students who have studied row and column operations for determinants (see Unit 7.3) may obtain this by simplifying the determinant first. (One way is to subtract the third column from the first column and then add the third row to the first row).

The eigenvalues are therefore $\lambda = -1$(repeated) and $\lambda = 8$.

(b) The eigenvectors

Case 1. $\lambda = 8$

We require to solve the homogeneous equations

$$\begin{aligned} -5x + 2y + 4z &= 0, \\ 2x - 8y + 2z &= 0, \\ 4x + 2y - 5z &= 0. \end{aligned}$$

Eliminating x from the second and third equations gives $18y - 9z = 0$.

Eliminating y from the second and third equations gives $18x - 18z = 0$.

Since z appears twice in the two statements, we may try letting $z = 1$ to give $y = \frac{1}{2}$ and $x = 1$.

Hence,

$$x : y : z = 1 : \frac{1}{2} : 1 = 2 : 1 : 2$$

The eigenvectors corresponding to $\lambda = 8$ are thus given by

$$\mathrm{X} = \alpha \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix},$$

where α is any **non-zero** scalar.

Case 2. $\lambda = -1$

This time, we require to solve the homogeneous equations

$$\begin{aligned} 4x + 2y + 4z &= 0, \\ 2x + y + 2z &= 0, \\ 4x + 2y + 4z &= 0, \end{aligned}$$

which, of course, are all the same equation, $2x + y + 2z = 0$.

Hence, two of the variables may be chosen at random (say $y = \beta$ and $z = \gamma$), then the third variable may be expressed in terms of them; (in this case $x = -\frac{1}{2}\beta - \gamma$).

However, a neater technique is to obtain, first, a pair of independent particular solutions by setting one pair of the variables at 1 and 0, in both orders. For example $y = 1$ and $z = 0$ gives $x = -\frac{1}{2}$ while $y = 0$ and $z = 1$ gives $x = -1$.

The general solution (in which y and z are randomly chosen as β and γ respectively) is then given by

$$\mathrm{X} = \beta \begin{bmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix},$$

where β and γ are **not both equal to zero** at the same time.

Notes:

(i) Similar results in Case 2 could be obtained by choosing **different** pairs of the three variables at random.

(ii) Other special cases arise if the three homogeneous equations reduce to a single equation in which one or even two of the variables is absent.

For example, if they reduced to $y = 0$, the corresponding eigenvectors could be given by

$$\mathrm{X} = \alpha \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

allowing $x = \alpha$ and $z = \gamma$ to be chosen at random, assuming that α and γ are not both zero simultaneously.

Alternatively, if the homogeneous equations reduced to $3x + 5z = 0$, then the corresponding eigenvectors could be given by

$$\mathrm{X} = \alpha \begin{bmatrix} 1 \\ 0 \\ -\frac{3}{5} \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

allowing $x = \alpha$ and $y = \beta$ to be chosen at random, assuming that α and β are not both zero simultaneously.

9.6.3 EXERCISES

Determine the eigenvalues and eigenvectors of the following matrices:

1.

$$\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}.$$

2.

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

3.

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

4.

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

5.

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

6.

$$\begin{bmatrix} -2 & -8 & -12 \\ 1 & 4 & 4 \\ 0 & 0 & 1 \end{bmatrix}.$$

7.

$$\begin{bmatrix} -3 & -9 & -12 \\ 1 & 3 & 4 \\ 0 & 0 & 1 \end{bmatrix}.$$

8.

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

9.

$$\begin{bmatrix} 2 & 0 & 0 \\ 3 & 2 & 0 \\ 4 & 0 & 2 \end{bmatrix}.$$

10.

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

9.6.4 ANSWERS TO EXERCISES

The eigenvectors will be written as X^T instead of X, in order to avoid unnecessary waste of space.

1. $\lambda = 0$, giving $X^T = \alpha[-3 \;\; 1]$,
   where α is any non-zero scalar;

   $\lambda = 7$, giving $X^T = \beta[1 \;\; 2]$,
   where β is any non-zero scalar.

2. $\lambda = 1$, giving $X^T = \alpha[1 \;\; -1 \;\; 0]$,
   where α is any non-zero scalar;

   $\lambda = 2$, giving $X^T = \beta[-2 \;\; 1 \;\; 2]$,
   where β is any non-zero scalar;

   $\lambda = 3$, giving $X^T = \gamma[-1 \;\; 1 \;\; 2]$,
   where γ is any non-zero scalar.

3. $\lambda = 1$, giving $X^T = \alpha[3 \;\; 2 \;\; 1]$,
   where α is any non-zero scalar;

   $\lambda = -1$, giving $X^T = \beta[1 \;\; 0 \;\; 1]$,
   where β is any non-zero scalar;

   $\lambda = 2$, giving $X^T = \gamma[1 \;\; 3 \;\; 1]$,
   where γ is any non-zero scalar.

4. $\lambda = 1$, giving $X^T = \alpha[1 \;\; 1 \;\; -1]$,
   where α is any non-zero scalar;

   $\lambda = 2$ (repeated), giving $X^T = \beta[2 \;\; 1 \;\; 0]$,
   where β is any non-zero scalar.

5. $\lambda = 1$ (repeated), giving $X^T = \alpha[1 \;\; 1 \;\; 1]$,
   where α is any non-zero scalar.

6. $\lambda = 0$, giving $X^T = \alpha[-4 \;\; 1 \;\; 0]$,
   where α is any non-zero scalar;

   $\lambda = 2$, giving $X^T = \beta[-2 \;\; 1 \;\; 0]$,
   where β is any non-zero scalar;

   $\lambda = -1$, giving $X^T = \gamma[28 \;\; -8 \;\; 3]$,
   where γ is any non-zero scalar.

7. $\lambda = 1$, giving $X^T = \alpha[-12 \;\; 4 \;\; 1]$,
   where α is any non-zero scalar;

   $\lambda = 0$ (repeated), giving $X^T = \beta[-3 \;\; 1 \;\; 0] + \gamma[-4 \;\; 0 \;\; 1]$,
   where β and γ are any scalar numbers which are not simultaneously zero.

8. $\lambda = 5$, giving $X^T = \alpha[1 \;\; 1 \;\; 1]$,
   where α is any non-zero scalar;

   $\lambda = 1$ (repeated), giving $X^T = \beta[-2 \;\; 1 \;\; 0] + \gamma[-1 \;\; 0 \;\; 1]$,
   where β and γ are any scalar numbers which are not simultaneously zero.

9. $\lambda = 2$ (repeated), giving $X^T = \beta[0 \;\; 1 \;\; 0] + \gamma[0 \;\; 0 \;\; 1]$
   where β and γ are any scalar numbers which are not simultaneously zero.

10. $\lambda = 2$, giving $X^T = \alpha[1 \;\; 1 \;\; 1]$,
    where α is any non-zero number;

    $\lambda = 3$, giving $X^T = \alpha[1 \;\; \frac{1}{2} \;\; 0] + \gamma[0 \;\; 0 \;\; 1]$,
    where α and γ are any scalar numbers which are not simultaneously zero.

"JUST THE MATHS"

UNIT NUMBER

9.7

MATRICES 7
(Linearly independent eigenvectors)
&
(Normalised eigenvectors)

by

A.J.Hobson

UNIT 9.7 - MATRICES 7

LINEARLY INDEPENDENT AND NORMALISED EIGENVECTORS

9.7.1 LINEARLY INDEPENDENT EIGENVECTORS

It is often useful to know if an $n \times n$ matrix, A, possesses a full set of n eigenvectors $X_1,\ X_2,\ X_3, \ldots, X_n$, which are **"linearly independent"**.

That is, they are **not** connected by any relationship of the form

$$a_1X_1 + a_2X_2 + a_3X_3 + \ldots \equiv 0,$$

where $a_1,\ a_2,\ a_3, \ldots$ are constants.

If the eigenvalues of A are distinct, it turns out that the eigenvectors are linearly independent; but, if any of the eigenvalues are repeated, further investigation may be necessary.

ILLUSTRATIONS

1. In Unit 9.6, it was shown that the matrix

$$A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

has eigenvalues $\lambda = 8$ and $\lambda = -1$(repeated), with corresponding eigenvectors,

$$\alpha \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix},$$

where α is any **non-zero** scalar and

$$\beta \begin{bmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix},$$

where β and γ are not both equal to zero at the same time.

The matrix, A, possesses a set of three linearly independent eigenvectors which may, conveniently, be chosen as

$$X_1 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \quad X_2 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \quad \text{and} \quad X_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

It is reasonably obvious that these are linearly independent, but a formal check would be to show that the matrix,

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

has rank 3.

2. It may be shown that the matrix,

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

has eigenvalues $\lambda = 2$ (repeated) and $\lambda = 1$, with corresponding eigenvectors,

$$\alpha \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \beta \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix},$$

where α and β are any non-zero numbers.

In this case, it is not possible to obtain a full set of three linearly independent eigenvectors.

9.7.2 NORMALISED EIGENVECTORS

It is sometimes convenient to use a set of **"normalised"** eigenvectors, which means that, for each eigenvector, the sum of the squares of its elements is equal to 1.

An eigenvector may be normalised if we multiply it by (plus or minus) the reciprocal of the square root of the sum of the squares of its elements.

ILLUSTRATIONS

1. A set of linearly independent normalised eigenvectors for the matrix,

$$\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix},$$

is

$$\frac{1}{3}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \quad \frac{1}{\sqrt{5}}\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \quad \text{and} \quad \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

2. A set of linearly independent normalised eigenvectors for the matrix,

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

is

$$\frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

9.7.3 EXERCISES

1. Determine the eigenvalues and a set of linearly independent normalised eigenvectors for the following 2×2 matrices:

$$\text{(a)} \quad \begin{bmatrix} 17 & -6 \\ 45 & -16 \end{bmatrix}, \quad \text{(b)} \quad \begin{bmatrix} 5 & -2 \\ 7 & -4 \end{bmatrix}, \quad \text{(c)} \quad \begin{bmatrix} 16 & -8 \\ 24 & -12 \end{bmatrix}.$$

2. Determine the eigenvalues and a set of linearly independent normalised eigenvectors for the following 3×3 matrices:

(a)

$$\begin{bmatrix} 2 & -3 & 1 \\ 3 & 1 & 3 \\ -5 & 2 & -4 \end{bmatrix};$$

(b)

$$\begin{bmatrix} 7 & 0 & 4 \\ 7 & 0 & 4 \\ 0 & 0 & 11 \end{bmatrix};$$

(c)

$$\begin{bmatrix} 2 & -2 & 2 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{bmatrix};$$

(d)

$$\begin{bmatrix} -2 & 0 & -14 \\ -7 & 5 & -14 \\ 0 & 0 & 5 \end{bmatrix};$$

(e)

$$\begin{bmatrix} 3 & 0 & 0 \\ 1 & 4 & 0 \\ 1 & 0 & 4 \end{bmatrix}.$$

9.7.4 ANSWERS TO EXERCISES

1. (a) The eigenvalues are 2 and -1.
   A set of linearly independent normalised eigenvectors is

$$\frac{1}{\sqrt{29}}\begin{bmatrix} 2 \\ 5 \end{bmatrix}, \quad \frac{1}{\sqrt{10}}\begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

   (b) The eigenvalues are 3 and -2.
   A set of linearly independent normalised eigenvectors is

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \frac{1}{\sqrt{53}}\begin{bmatrix} 2 \\ 7 \end{bmatrix}.$$

   (c) The eigenvalues are 4 and 0.
   A set if linearly independent normalised eigenvectors is

$$\frac{1}{\sqrt{13}}\begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

2. (a) The eigenvalues are 1, 0 and -2.
A set of linearly independent normalised eigenvectors is

$$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\-1\end{bmatrix}, \quad \frac{1}{\sqrt{230}}\begin{bmatrix}10\\3\\-11\end{bmatrix} \quad \text{and} \quad \frac{1}{\sqrt{74}}\begin{bmatrix}4\\3\\-7\end{bmatrix}.$$

(b) The eigenvalues are 11, 7 and 0.
A set of linearly independent normalised eigenvectors are

$$\frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}, \quad \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}0\\1\\0\end{bmatrix}.$$

(c) The eigenvalues are 2 (repeated) and -2.
A set of linearly independent normalised eigenvectors is

$$\frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\1\end{bmatrix}, \quad \text{and} \quad \frac{1}{\sqrt{66}}\begin{bmatrix}4\\1\\7\end{bmatrix}.$$

(d) The eigenvalues are 5 (repeated) and -2.
A set of linearly independent normalised eigenvectors is

$$\frac{1}{\sqrt{6}}\begin{bmatrix}-2\\1\\1\end{bmatrix}, \quad \frac{1}{\sqrt{6}}\begin{bmatrix}-2\\0\\1\end{bmatrix} \quad \text{and} \quad \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\end{bmatrix}.$$

(e) The eigenvalues are 4 (repeated) and 3.
A set of linearly independent normalised eigenvectors is

$$\begin{bmatrix}0\\1\\0\end{bmatrix}, \quad \begin{bmatrix}0\\0\\1\end{bmatrix} \quad \text{and} \quad \frac{1}{\sqrt{3}}\begin{bmatrix}1\\-1\\-1\end{bmatrix}.$$

"JUST THE MATHS"

UNIT NUMBER

9.8

MATRICES 8
(Characteristic properties)
&
(Similarity transformations)

by

A.J.Hobson

UNIT 9.8 - MATRICES 8

CHARACTERISTIC PROPERTIES AND SIMILARITY TRANSFORMATIONS

9.8.1 PROPERTIES OF EIGENVALUES AND EIGENVECTORS

We list, here, a number of standard properties, together with either their proofs or an illustration of their proofs.

(i) The eigenvalues of a matrix are the same as those of its transpose.

Proof:

Given a square matrix, A, the eigenvalues of A^{T} are the solutions of the equation

$$|\mathrm{A}^{\mathrm{T}} - \lambda\mathrm{I}| = 0.$$

But, since I is a symmetric matrix, this is equivalent to

$$|(\mathrm{A} - \lambda\mathrm{I})^{\mathrm{T}}| = 0.$$

The result follows since a determinant is unchanged in value when it is transposed.

(ii) The eigenvalues of the multiplicative inverse of a matrix are the reciprocals of the eigenvalues of the matrix itself.

Proof:

If λ is any eigenvalue of a square matrix, A, then

$$\mathrm{AX} = \lambda\mathrm{X},$$

for some column vector, X.

Premultiplying this relationship by A^{-1}, we obtain

$$\mathrm{A}^{-1}\mathrm{AX} = \mathrm{A}^{-1}(\lambda\mathrm{X}) = \lambda(\mathrm{A}^{-1}\mathrm{X}).$$

Thus,

$$A^{-1}X = \frac{1}{\lambda}X.$$

(iii) The eigenvectors of a matrix and its multiplicative inverse are the same.

Proof:

This follows from the proof of **(ii)**, since

$$A^{-1}X = \frac{1}{\lambda}X$$

implies that X is an eigenvector of A^{-1}.

(iv) If a matrix is multiplied by a single number, the eigenvalues are multiplied by that number, but the eigenvectors remain the same.

Proof:

If A is multiplied by α, we may write the equation $AX = \lambda X$ in the form $\alpha AX = \alpha\lambda X$.

Thus, αA has eigenvalues, $\alpha\lambda$, and eigenvectors, X.

(v) If $\lambda_1, \quad \lambda_2, \quad \lambda_3, \ . \ .$ are the eigenvalues of the matrix A and n is a positive integer, then $\lambda_1^n, \quad \lambda_2^n, \quad \lambda_3^n, \ . \ .$ are the eigenvalues of A^n.

Proof:

If λ denotes any one of the eigenvalues of the matrix, A, then $AX = \lambda X$.

Premultiplying both sides by A, we obtain $A^2X = A\lambda X = \lambda AX = \lambda^2 X$

Hence, λ^2 is an eigenvlaue of A^2.

Similarly, $A^3X = \lambda^3 X$, and so on.

(vi) If $\lambda_1, \ \lambda_2, \ \lambda_3, \ \ . \ . \ .$ are the eigenvalues of the $n \times n$ matrix A, I is the $n \times n$ multiplicative identity matrix and k is a single number, then the eigenvalues of the matrix $A + kI$ are $\lambda_1 + k, \ \lambda_2 + k, \ \lambda_3 + k, \ . \ . \ ..$

Proof:

If λ is any eigenvalue of A, then $AX = \lambda X$.

Hence,

$$(A + kI)X = AX + kX = \lambda X + kX = (\lambda + k)X.$$

(vii) A matrix is singular ($|A| = 0$) if and only if at least one eigenvalue is equal to zero.

Proof:

(a) If X is an eigenvector corresponding to an eigenvalue, $\lambda = 0$, then $AX = \lambda X = [0]$.

From the theory of homogeneous linear equations (see Unit 7.4), it follows that $|A| = 0$.

(b) Conversely, if $|A| = 0$, the homogeneous system $AX = [0]$ has a solution for X other than $X = [0]$. Hence, at least one eigenvalue must be zero.

(viii) If A is an orthogonal matrix ($AA^T = I$), then every eigenvalue is either $+1$ or -1.

Proof:

The statement $AA^T = I$ can be written $A^{-1} = A^T$ so that, by **(i)** and **(ii)**, the eigenvalues of A are equal to their own reciprocals.

That is, they must have values $+1$ or -1.

(ix) If the elements of a matrix below the leading diagonal or the elements above the leading diagonal are all equal zero, then the eigenvalues are equal to the diagonal elements.

ILLUSTRATION

An **"upper-triangular matrix"**, A, of order 3×3, has the form

$$A = \begin{bmatrix} a_1 & b_1 & c_1 \\ 0 & b_2 & c_2 \\ 0 & 0 & c_3 \end{bmatrix}.$$

The characteristic equation is given by

$$0 = |\mathrm{A} - \lambda \mathrm{I}| = \begin{vmatrix} a_1 - \lambda & b_1 & c_1 \\ 0 & b_2 - \lambda & c_2 \\ 0 & 0 & c_3 - \lambda \end{vmatrix} = (a_1 - \lambda)(b_2 - \lambda)(c_3 - \lambda).$$

Hence, $\lambda = a_1,\ b_2$ or c_3.

A similar proof holds for a **"lower-triangular matrix"**.

Note:
A special case of both a lower-triangular matrix and an upper-triangular matrix is a diagonal matrix.

(x) The sum of the eigenvalues of a matrix is equal to the trace of the matrix (the sum of the diagonal elements) and the product of the eigenvalues is equal to the determinant of the matrix.

ILLUSTRATION

We consider the case of a 2×2 matrix, A, given by

$$\mathrm{A} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}.$$

The characteristic equation is

$$0 = \begin{vmatrix} a_1 - \lambda & b_1 \\ a_2 & b_2 - \lambda \end{vmatrix} = \lambda^2 - (a_1 + b_2)\lambda + (a_1 b_2 - a_2 b_1).$$

But, for any quadratic equation, $a\lambda^2 + b\lambda + c = 0$, the sum of the solutions is equal to $-b/a$ and the product of the solutions is equal to c/a.

In this case, therefore, the sum of the solutions is $a_1 + b_2$, while the product of the solutions is $a_1 b_2 - a_2 b_1$.

9.8.2 SIMILAR MATRICES

In the previous section, a matrix and its transpose illustrated how two matrices can have the same eigenvalues. In this section, we deal with a more general case of this occurrence.

DEFINITION

Two matrices, A and B, are said to be **"similar"** if

$$B = P^{-1}AP,$$

for some non-singular matrix, P.

Notes:

(i) P is certainly square, so that A and B must also be square and of the same order as P.

(ii) The relationship $B = P^{-1}AP$ is regarded as a **"transformation"** of the matrix, A, into the matrix, B.

(iii) A relationship of the form $B = QAQ^{-1}$ may also be regarded as a similarity transformation on A, since Q is the multiplicative inverse of Q^{-1}.

THEOREM

Two similar matrices, A and B, have the same eigenvalues. Furthermore, if the similarity transformation from A to B is $B = P^{-1}AP$, then the eigenvectors, X and Y, of A and B respectively are related by the equation

$$Y = P^{-1}X.$$

Proof:

The eigenvalues, λ, and the eigenvectors, X, of A satisfy the relationship $AX = \lambda X$.

Hence,

$$P^{-1}AX = \lambda P^{-1}X.$$

Secondly, using the fact that $PP^{-1} = I$, we have

$$P^{-1}APP^{-1}X = \lambda P^{-1}X,$$

which may be written

$$(P^{-1}AP)(P^{-1}X) = \lambda(P^{-1}X)$$

or

$$BY = \lambda Y,$$

where $B = P^{-1}AP$ and $Y = P^{-1}X$.

This shows that the eigenvalues of A are also the eigenvalues of B and that the eigenvectors of B are of the form $P^{-1}X$.

Reminders

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

and, in general, for a square matrix M,

$$M^{-1} = \tfrac{1}{|M|} \times \text{ the transpose of the cofactor matrix.}$$

9.8.3 EXERCISES

1. For the matrix,

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

   and its multiplicative inverse, A^{-1}, determine the eigenvalues and a set of corresponding linearly independent normalised eigenvectors.

2. State the eigenvalues for the upper-triangular matrix

$$\begin{bmatrix} 2 & -4 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & -1 \end{bmatrix}$$

and, hence, obtain a set of linearly independent normalised eigenvectors for the matrix.

3. State the eigenvalues of the lower-triangular matrix

$$\begin{bmatrix} 6 & 0 & 0 \\ 3 & 0 & 0 \\ 2 & 1 & -10 \end{bmatrix}$$

and, hence, obtain a set of linearly independent normalised eigenvectors for the matrix.

4. Determine the eigenvalues and a set of corresponding linearly independent eigenvectors for the matrix $B = P^{-1}AP$, where

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \quad \text{and} \quad P = \begin{bmatrix} 4 & 7 \\ 1 & 2 \end{bmatrix}.$$

5. Determine the eigenvalues and a set of corresponding linearly independent eigenvectors for the matrix $B = P^{-1}AP$, where

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 4 & 0 \\ -2 & -1 & 1 \end{bmatrix} \quad \text{and} \quad P = \begin{bmatrix} 2 & -1 & 3 \\ 0 & -2 & 4 \\ 5 & 1 & 6 \end{bmatrix}.$$

6. Show that the matrix

$$C = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{\sqrt{2}}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \end{bmatrix}$$

is orthogonal and verify that its eigenvalues are either 1 or -1.

9.8.4 ANSWERS TO EXERCISES

1. The eigenvalues of A are 3, 2 and 1 with corresponding normalised eigenvectors,

$$\frac{1}{\sqrt{5}}\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \quad \frac{1}{\sqrt{3}}\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

The eigenvalues of A^{-1} are $\frac{1}{3}$, $\frac{1}{2}$ and 1, with corresponding normalised eigenvectors the same as for A.

2. The eigenvalues are 3, 2 and -1, with corresponding linearly independent normalised eigenvectors,

$$\frac{1}{\sqrt{17}}\begin{bmatrix} 4 \\ -1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ -2 \end{bmatrix}.$$

3. The eigenvalues are 6, 0 and -10, with corresponding linearly independent normalised eigenvectors,

$$\frac{1}{\sqrt{1305}}\begin{bmatrix} 32 \\ 16 \\ 5 \end{bmatrix}, \quad \frac{1}{\sqrt{101}}\begin{bmatrix} 0 \\ 10 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

4. The eigenvalues of B are the same as those of A, namely 0 and 7.

$$\text{Also,} \quad \mathrm{P}^{-1} = \begin{bmatrix} 2 & -7 \\ -1 & 4 \end{bmatrix}.$$

Hence, a set of linearly independent eigenvectors for B is

$$\begin{bmatrix} -13 \\ 7 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -12 \\ 7 \end{bmatrix}.$$

5. The eigenvalues of B are the same as those of A which, from question 1, are 3, 2 and 1. Also,

$$\mathrm{P}^{-1} = -\frac{1}{22}\begin{bmatrix} -16 & 9 & 2 \\ 20 & -3 & -8 \\ 10 & -7 & -4 \end{bmatrix},$$

so that a set of linearly independent eigenvectors for B are

$$\begin{bmatrix} -17 \\ 13 \\ 12 \end{bmatrix}, \quad \begin{bmatrix} -27 \\ 31 \\ 21 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1 \\ 4 \\ 2 \end{bmatrix}.$$

6. $\mathrm{CC}^{\mathrm{T}} = \mathrm{I}$ and the eigenvalues are 1 (repeated) and -1.

"JUST THE MATHS"

UNIT NUMBER

9.9

MATRICES 9
(Modal & spectral matrices)

by

A.J.Hobson

UNIT 9.9 - MATRICES 9

MODAL AND SPECTRAL MATRICES

9.9.1 ASSUMPTIONS AND DEFINITIONS

For convenience, we shall make, here, the following assumptions:

(a) The n eigenvalues, $\lambda_1,\ \lambda_2,\ \lambda_3, \ldots, \lambda_n$, of an $n \times n$ matrix, A, are arranged in order of decreasing value.

(b) Corresponding to $\lambda_1,\ \lambda_2,\ \lambda_3, \ldots, \lambda_n$, respectively, A possesses a full set of eigenvectors $X_1,\ X_2,\ X_3, \ldots, X_n$, which are linearly independent.

If two eigenvalues coincide, the order of writing down the corresponding pair of eigenvectors will be immaterial.

DEFINITION 1

The square matrix obtained by using as its columns any set of linearly independent eigenvectors of a matrix A is called a **"modal matrix"** of A, and may be denoted by M.

Notes:

(i) There are infinitely many modal matrices for a given matrix, A, since any multiple of an eigenvector is also an eigenvector.

(ii) It is sometimes convenient to use a set of normalised eigenvectors.

When using normalised eigenvectors, the modal matrix may be denoted by N and, for an $n \times n$ matrix, A, there are 2^n possibilities for N since each of the n columns has two possibilities.

DEFINITION 2

If $\lambda_1,\ \lambda_2,\ \lambda_3,\ \ldots, \lambda_n$ are the eigenvalues of an $n \times n$ matrix, A, then the diagonal matrix,

$$\begin{bmatrix} \lambda_1 & 0 & 0 & . & . & 0 \\ 0 & \lambda_2 & 0 & . & . & 0 \\ 0 & 0 & \lambda_3 & . & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & \lambda_n \end{bmatrix}$$

is called the **"spectral matrix"** of A, and may be denoted by S.

EXAMPLE

For the matrix,

$$A = \begin{bmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix},$$

determine a modal matrix, a modal matrix of normalised eigenvectors and the spectral matrix.

Solution

The characteristic equation is

$$\begin{vmatrix} 1-\lambda & 1 & -2 \\ -1 & 2-\lambda & 1 \\ 0 & 1 & -1-\lambda \end{vmatrix} = 0,$$

which may be shown to give

$$-(1+\lambda)(1-\lambda)(2-\lambda) = 0.$$

Hence, the eigenvalues are $\lambda_1 = 2,\ \lambda_2 = 1$ and $\lambda_3 = -1$ in order of decreasing value.

Case 1. $\lambda = 2$

We solve the simultaneous equations

$$\begin{aligned} -x + y - 2z &= 0, \\ -x + 0y + z &= 0, \\ 0x + y - 3z &= 0, \end{aligned}$$

which give $x : y : z = 1 : 3 : 1$

Case 2. $\lambda = 1$

We solve the simultaneous equations

$$\begin{aligned} 0x + y - 2z &= 0, \\ -x + y + z &= 0, \\ 0x + y - 2z &= 0, \end{aligned}$$

which give $x : y : z = 3 : 2 : 1$

Case 3. $\lambda = -1$

We solve the simultaneous equations

$$\begin{aligned} 2x + y - 2z &= 0, \\ -x + 3y + z &= 0, \\ 0x + y + 0z &= 0, \end{aligned}$$

which give $x : y : z = 1 : 0 : 1$

A modal matrix for A may therefore be given by

$$\mathrm{M} = \begin{bmatrix} 1 & 3 & 1 \\ 3 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

A modal matrix of normalised eigenvectors may be given by

$$\mathrm{N} = \begin{bmatrix} \frac{1}{\sqrt{11}} & \frac{3}{\sqrt{14}} & \frac{1}{\sqrt{2}} \\ \frac{3}{\sqrt{11}} & \frac{2}{\sqrt{14}} & 0 \\ \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{14}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

The spectral matrix is given by

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

9.9.2 DIAGONALISATION OF A MATRIX

Since the eigenvalues of a diagonal matrix are equal to its diagonal elements, it is clear that a matrix, A, and its spectral matrix, S, have the same eigenvalues.

From the Theorem in Unit 9.8, therefore, it seems reasonable that A and S could be similar matrices; and this is the content of the following result which will be illustrated rather than proven.

THEOREM

The matrix, A, is similar to its spectral matrix, S, the similarity transformation being

$$M^{-1}AM = S,$$

where M is a modal matrix for A.

ILLUSTRATION:

Suppose that $X_1,\ X_2$ and X_3 are linearly independent eigenvectors of a 3×3 matrix, A, corresponding to eigenvalues $\lambda_1,\ \lambda_2$ and λ_3, respectively.

Then,

$$AX_1 = \lambda_1 X_1, \quad AX_2 = \lambda_2 X_2, \quad \text{and} \quad AX_3 = \lambda_3 X_3.$$

Also,

$$M = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix}.$$

If M is premultiplied by A, we obtain a 3×3 matrix whose columns are $AX_1,\ AX_2,$ and AX_3.

That is,

$$AM = [\,AX_1 \quad AX_2 \quad AX_3\,] = [\,\lambda_1 X_1 \quad \lambda_2 X_2 \quad \lambda_3 X_3\,]$$

or

$$AM = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix} . \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = MS.$$

We conclude that

$$M^{-1}AM = S.$$

Notes:

(i) M^{-1} exists only because X_1, X_2 and X_3 are linearly independent.

(ii) The similarity transformation in the above theorem reduces the matrix, A, to **"diagonal form"** or **"canonical form"** and the process is often referred to as the **"diagonalisation"** of the matrix, A.

EXAMPLE

Verify the above Theorem for the matrix

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

Solution

From an earlier example, a modal matrix for A may be given by

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

The spectral matrix is given by

$$\mathrm{S} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

It may be shown that

$$\mathrm{M}^{-1} = \frac{1}{6}\begin{bmatrix} -2 & 2 & 2 \\ 3 & 0 & -3 \\ -1 & -2 & 7 \end{bmatrix}$$

and, hence,

$$\mathrm{M}^{-1}\mathrm{AM} = \frac{1}{6}\begin{bmatrix} -2 & 2 & 2 \\ 3 & 0 & -3 \\ -1 & -2 & 7 \end{bmatrix} . \begin{bmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix} . \begin{bmatrix} 1 & 3 & 1 \\ 3 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

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

$$= \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \mathrm{S}.$$

9.9.3 EXERCISES

1. Determine a modal matrix, M, of linearly independent eigenvectors for the matrix

$$A = \begin{bmatrix} -1 & 6 & -12 \\ 0 & -13 & 30 \\ 0 & -9 & 20 \end{bmatrix}.$$

   Verify that $M^{-1}AM = S$, where S is the spectral matrix of A.

2. Determine a modal matrix, M, of linearly independent eigenvectors for the matrix

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

   Verify that $M^{-1}AM = S$, where S is the spectral matrix of B.

3. Determine a modal matrix, N, of linearly independent normalised eigenvectors for the matrix

$$C = \begin{bmatrix} 2 & -2 & 0 \\ 0 & 4 & 0 \\ 2 & 5 & 1 \end{bmatrix}.$$

   Verify that $N^{-1}AN = S$, where S is the spectral matrix of C.

4. Show that the following matrices are not similar to a diagonal matrix:

$$\text{(a)} \quad \begin{bmatrix} 1 & 2 & 2 \\ 0 & 2 & 1 \\ -1 & 2 & 2 \end{bmatrix}, \qquad \text{(b)} \quad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{bmatrix}.$$

9.9.4 ANSWERS TO EXERCISES

1. The eigenvalues are 5,2 and -1, which gives

$$\mathrm{M} = \begin{bmatrix} -1 & 0 & 1 \\ 5 & 2 & 0 \\ 3 & 1 & 0 \end{bmatrix}.$$

2. The eigenvalues are 2, 1 and -1, which gives

$$\mathrm{M} = \begin{bmatrix} 1 & 3 & 1 \\ 3 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

3. The eigenvalues are 4, 2 and 1, which gives

$$\mathrm{N} = \begin{bmatrix} -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{5}} & 0 \\ \frac{1}{\sqrt{3}} & 0 & 0 \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{5}} & 1 \end{bmatrix}.$$

4. (a) The eigenvalues are 2 (repeated) and 1 but there are only two linearly independent eigenvectors, namely

$$\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

(b) There is only one eigenvalue, 1 (repeated), and only one linearly independent eigenvector, namely

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

"JUST THE MATHS"

UNIT NUMBER

9.10

MATRICES 10
(Symmetric matrices & quadratic forms)

by

A.J.Hobson

UNIT 9.10 - MATRICES 10

SYMMETRIC MATRICES AND QUADRATIC FORMS

9.10.1 SYMMETRIC MATRICES

The definition of a symmetric matrix was introduced in Unit 9.1 and matrices of this type have certain special properties with regard to eigenvalues and eigenvectors. We list them as follows:

(i) All of the eigenvalues of a symmetric matrix are real and, hence, so are the eigenvectors.

(ii) A symmetric matrix of order $n \times n$ always has n linearly independent eigenvectors.

(iii) For a symmetric matrix, suppose that X_i and X_j are linearly independent eigenvectors associated with different eigenvectors; then

$$\mathrm{X}_i\mathrm{X}_j^{\mathrm{T}} \equiv x_ix_j + y_iy_j + z_iz_j = 0.$$

We say that X_i and X_j are **"mutually orthogonal"**.

If a symmetric matrix has any repeated eigenvalues, it is still possible to determine a full set of mutually orthogonal eigenvectors, but not every full set of eigenvectors will have the orthogonality property.

(iv) A symmetric matrix always has a modal matrix whose columns are mutually orthogonal. When the eigenvalues are distinct, this is true for every modal matrix.

(v) A modal matrix, N, of normalised eigenvectors is an orthogonal matrix.

ILLUSTRATIONS

1. If N is of order 3×3, we have

$$\mathrm{N}^{\mathrm{T}}.N = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix} . \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

2. It was shown in Unit 9.6 that the matrix

$$A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

has eigenvalues $\lambda = 8$, and $\lambda = -1$ (repeated), with associated eigenvectors

$$\alpha \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \text{ and } \beta \begin{bmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \equiv \begin{bmatrix} -\frac{1}{2}\beta - \gamma \\ \beta \\ \gamma \end{bmatrix}.$$

A set of **linearly independent** eigenvectors may therefore be given by

$$X_1 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \quad X_2 = \begin{bmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad X_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Clearly, X_1 is orthogonal to X_2 and X_3, but X_2 and X_3 are not orthogonal to each other. However, we may find β and γ such that

$$\beta \begin{bmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \text{ is orthogonal to } \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

We simply require that

$$\frac{1}{2}\beta + 2\gamma = 0$$

or

$$\beta + 4\gamma = 0;$$

and this will be so, for example, when $\beta = 4$ and $\gamma = -1$.

A new set of linearly independent mutually orthogonal eigenvectors can thus be given by

$$X_1 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \quad X_2 = \begin{bmatrix} -1 \\ 4 \\ -1 \end{bmatrix}, \quad \text{and} \quad X_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

9.10.2 QUADRATIC FORMS

An algebraic expression of the form

$$ax^2 + by^2 + cz^2 + 2fyz + 2yzx + 2hxy$$

is called a **"quadratic form"**.

In matrix notation, it may be written as

$$[x \quad y \quad z] \begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \equiv X^T AX,$$

and we note that the matrix A is symmetric.

In the scientific applications of quadratic forms, it is desirable to know whether such a form is

(a) always positive,

(b) always negative,

(c) both positive and negative.

It may be shown that, if we change to new variables, (u, v, w), using a linear transformation

$$X = PU,$$

where P is some non-singular matrix, then the new quadratic form has the same properties as the original, concerning its sign.

We now show that a good choice for P is a modal matrix, N, of normalised, linearly independent, mutually orthogonal eigenvectors for A.

Putting X = NU, the expression X^TAX becomes U^TN^TANX.

But, since N is orthogonal when A is symmetric, $N^T = N^{-1}$ and, hence, N^TAN is the spectral matrix, S, for A.

The new quadratic form is therefore

$$U^TSU \equiv [\,u \quad v \quad w\,] . \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} . \begin{bmatrix} u \\ v \\ w \end{bmatrix} \equiv \lambda_1 u^2 + \lambda_2 v^2 + \lambda_3 w^2.$$

Clearly, if all of the eigenvalues are positive, then the new quadratic form is always positive; and, if all of the eigenvalues are negative, then the new quadratic form is always negative.

The new quadratic form is called the **"canonical form under similarity"** of the original quadratic form.

9.10.3 EXERCISES

1. For the following symmetric matrices, determine a set of three linearly independent and mutually orthogonal eigenvectors:

$$\text{(a)} \quad \begin{bmatrix} 2 & 0 & 0 \\ 0 & 10 & 6 \\ 0 & 6 & 5 \end{bmatrix}, \quad \text{(b)} \quad \begin{bmatrix} 2 & 2 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

2. Repeat the previous question for the following symmetric matrices:

$$\text{(a)} \quad \begin{bmatrix} 3 & 3 & 3\sqrt{2} \\ 3 & 3 & 3\sqrt{2} \\ 3\sqrt{2} & 3\sqrt{2} & 6 \end{bmatrix}, \quad \text{(b)} \quad \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & -2 \end{bmatrix}.$$

3. Using the results of question 1, show that the following quadratic forms are always positive:

   (a)

$$2x^2 + 10y^2 + 5z^2 + 12yz;$$

   (b)

$$2x^2 + 5y^2 + 3z^2 + 4xy.$$

4. Using the results of question 2(b), obtain the matrix, P, of the orthogonal transformation, X = PU, which transforms the quadratic function

$$2x^2 + y^2 - 2z^2 + 4xz$$

into the quadratic function

$$2u^2 + 2v^2 - 3w^2.$$

State whether or the not the original quadratic form is always positive.

9.10.4 ANSWERS TO EXERCISES

1. (a) The eigenvalues are 14, 2 and 1 and a set of linearly independent mutually orthogonal eigenvectors is

$$\begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ -2 \\ 3 \end{bmatrix}.$$

(b) The eigenvalues are 6, 3 and 1 and a set of linearly independent mutually orthogonal eigenvectors is

$$\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}.$$

2. (a) The eigenvalues are 12 and 0 (repeated) and set of linearly independent mutually orthogonal eigenvectors is

$$\begin{bmatrix} 1 \\ 1 \\ \sqrt{2} \end{bmatrix}, \quad \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ 1 \\ -\sqrt{2} \end{bmatrix}.$$

(b) The eigenvalues are 2 and -3 (repeated) and a set of linearly independent mutually orthogonal eigenvectors is

$$\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}.$$

3. The eigenvalues are all positive and hence the quadratic forms are always positive.
4. $$\mathrm{P} = \begin{bmatrix} 0 & 1 & 0 \\ \frac{2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & 0 & -\frac{2}{\sqrt{5}} \end{bmatrix}.$$

   The original quadratic form may take both positive and negative values since the associated eigenvalues are not all positive.

“JUST THE MATHS”

UNIT NUMBER

10.1

DIFFERENTIATION 1
(Functions and limits)

by

A.J.Hobson

UNIT 10.1 - DIFFERENTIATION 1

10.1.1 FUNCTIONAL NOTATION

Introduction

If a variable quantity, y, depends for its values on another variable quantity, x, we say that **"y is a function of x"** and we write, in general:

$$y = f(x),$$

which is pronounced "y equals f of x".

Notes:

(i) y is called the **"dependent variable"** and x is called the **"independent variable"**. The choice of value for x will be arbitrary, within certain possible constraints, but, once it **is** chosen, the value of y is then fixed.

(ii) The use of the letter f in $f(x)$ is logical because it stands for the word "function"; but once the format of the notation is understood, we can use other letters as appropriate. For example:
the statement $P = P(T)$ could be used to indicate that a pressure, P, is a function of absolute temperature, T;
the statement $i = i(t)$ could be used to indicate that an electric current, i, is a function of time t;
the original statement could have been written $y = y(x)$ without using f at all.

The general format of functional notation may be described as follows:

DEPENDENT VARIABLE =
DEPENDENT VARIABLE(INDEPENDENT VARIABLE)

10.1.2 NUMERICAL EVALUATION OF FUNCTIONS

If α is a number, then $f(\alpha)$ denotes the value of the function $f(x)$ when $x = \alpha$ is substituted into it.

For example, if

$$f(x) \equiv 4\sin 3x,$$

then,

$$f\left(\frac{\pi}{4}\right) = 4\sin\frac{3\pi}{4} = 4 \times \frac{1}{\sqrt{2}} \cong 2.828$$

10.1.3 FUNCTIONS OF A LINEAR FUNCTION

The notation

$$f(ax+b),$$

where a and b are constants, implies a **known** function, $f(x)$, in which x has been replaced by the linear function $ax+b$.

For example, if

$$f(x) \equiv 3x^2 - 7x + 4,$$

then,

$$f(5x-1) \equiv 3(5x-1)^2 - 7(5x-1) + 4;$$

but, in the applications of this kind of notation, it usually best to leave the expression in the bracketed form rather than to expand out the brackets and so lose any obvious connection between $f(x)$ and $f(ax+b)$.

10.1.4 COMPOSITE FUNCTIONS (or Functions of a Function) IN GENERAL

The symbol

$$f[g(x)]$$

implies a **known** function, $f(x)$, in which x has been replaced by **another known** function, $g(x)$.

For example, if

$$f(x) \equiv x^2 + 2x - 5$$

and

$$g(x) \equiv \sin x,$$

then,

$$f[g(x)] \equiv \sin^2 x + 2\sin x - 5;$$

but we can observe also that

$$g[f(x)] \equiv \sin(x^2 + 2x - 5),$$

which is not identical to the first result. Hence, in general,

$$f[g(x)] \not\equiv g[f(x)].$$

There are some exceptions to this, however, as in the case when

$$f(x) \equiv e^x \quad \text{and} \quad g(x) \equiv \log_e x,$$

whereupon we obtain

$$f[g(x)] \equiv e^{\log_e x} \equiv x$$

and

$$g[f(x)] \equiv \log_e (e^x) \equiv x.$$

The functions $\log_e x$ and e^x are said to be **"inverses"** of each other.

10.1.5 INDETERMINATE FORMS

Certain fractional expressions involving functions can become problematic if the values of the variable being substituted into them reduce them to either of the forms

$$\frac{0}{0} \text{ or } \frac{\infty}{\infty}.$$

Both of these forms are meaningless or **"indeterminate"** and need to be dealt with using a concept called **"limiting values"**.

(a) The Indeterminate Form $\frac{0}{0}$

Suppose the fractional expression

$$\frac{f(x)}{g(x)}$$

is such that both $f(x)$ **and** $g(x)$ take the value zero when $x = \alpha$; that is, $f(\alpha) = 0$ and $g(\alpha) = 0$. It is impossible, therefore, to evaluate the fraction when $x = \alpha$; but we may consider its values as x becomes increasingly close to α with out actually reaching it. The standard terminology is to say that "x **tends to** α", written $x \to \alpha$, for short.

We note that, by the **Factor Theorem**, discussed in Unit 1.8, $(x - \alpha)$ must be a factor of both numerator and denominator; and it turns out that, by cancelling this common factor (which is allowed if x is not going to reach α) we can assign a value to $\frac{f(x)}{g(x)}$ called a limiting value. It still will not be the value of this fraction **at** $x = \alpha$, but represents the value it approaches as x **tends** to α. The result is denoted by

$$\lim_{x \to \alpha} \frac{f(x)}{g(x)}.$$

EXAMPLE
Calculate

$$\lim_{x \to 1} \frac{x - 1}{x^2 + 2x - 3}.$$

Solution

First we factorise the denominator, knowing already that one of its factors must be $x - 1$ because it takes the value zero when $x = 1$.

The result is therefore

$$\lim_{x\to 1} \frac{x-1}{(x-1)(x+3)}$$

$$= \lim_{x\to 1} \frac{1}{x+3}.$$

What we now need to establish is the fixed value which this new fraction approaches as x becomes increasingly close to 1. But since there are no longer any problems with indeterminate forms, we do in fact simply substitute $x = 1$, obtaining the number $\frac{1}{4}$.

Hence,

$$\lim_{x\to 1} \frac{x-1}{x^2+2x-3} = \frac{1}{4}.$$

(b) The Indeterminate Form $\frac{\infty}{\infty}$

This kind of indeterminate form is usually encountered when the value of x itself becomes infinite, either positively or negatively. The object is to evaluate either

$$\lim_{x\to\infty} \frac{f(x)}{g(x)}$$

or

$$\lim_{x\to-\infty} \frac{f(x)}{g(x)}.$$

EXAMPLE
Calculate

$$\lim_{x\to\infty} \frac{2x^2+3x-1}{7x^2-2x+5}.$$

Solution

There is no factorising to do in this type of exercise; we simply divide the numerator and the denominator by the highest power of x appearing, then allow x to become infinite.

The result is therefore

$$\lim_{x\to\infty} \frac{2+\frac{3}{x}-\frac{1}{x^2}}{7-\frac{2}{x}+\frac{5}{x^2}} = \frac{2}{7}.$$

Note:
In the case of the ratio of two polynomials of equal degree, the limiting value as $x \to \pm\infty$ will always be the ratio of the leading coefficients of x. The same principle can be applied to the ratio of two polynomials of unequal degree if we insert zero coefficients in appropriate places to consider them as being of equal degree. The results then obtained will be either zero or infinity.

ILLUSTRATION

$$\lim_{x\to\infty} \frac{5x+11}{3x^2-4x+1} = \lim_{x\to\infty} \frac{0x^2+5x+11}{3x^2-4x+1} = \frac{0}{3} = 0.$$

A Useful Standard Limit

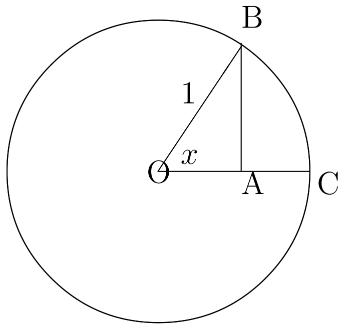
In Unit 3.3, it is shown that, for very small values of x in radians, $\sin x \simeq x$.

This suggests that

$$\lim_{x\to 0} \frac{\sin x}{x} = 1,$$

though we may not actually use the result from Unit 3.3 to **prove** the validity of this new limiting value. We shall see later that "$\sin x \simeq x$ for small x" is developed from a calculus technique which **assumes** that $\frac{\sin x}{x} \to 1$ as $x \to 0$; so, we would be using the result to prove itself !

An alternative, non-rigorous, proof is to consider the following diagram in which the angle x is situated at the centre of a circle with radius 1:



In the diagram, the length of line AB $= \sin x$ and the length of arc BC $= x$. Furthermore, as x decreases almost to zero, the two lengths become closer and closer to each other in value. That is,

$$\sin x \to x \quad \text{as} \quad x \to 0$$

or

$$\lim_{x \to 0} \frac{\sin x}{x} = 1$$

10.1.6 EVEN AND ODD FUNCTIONS

It is easy to see that any **even** power of x will be unchanged in value if x is replaced by $-x$. In a similar way, any **odd** power of x will be unchanged in numerical value, though altered in sign, if x is replaced by $-x$. These two powers of x are examples of an **"even function"** and an **"odd function"** respectively; but the true definition includes a wider range of functions as follows:

DEFINITION
A function $f(x)$ is said to be **"even"** if it satisfies the identity

$$f(-x) \equiv f(x).$$

ILLUSTRATIONS: x^2, $2x^6 - 4x^2 + 5$, $\cos x$.

DEFINITION
A function $f(x)$ is aid to be **"odd"** if it satisfies the identity

$$f(-x) \equiv -f(x).$$

ILLUSTRATIONS x^3, $x^5 - 3x^3 + 2x$, $\sin x$.

Note:
It is not necessary for every function to be either even or odd. For example, the function $x + 3$ is neither even nor odd.

EXAMPLE

Express an arbitrary function, $f(x)$, as the sum of an even function and an odd function.

Solution

We may write

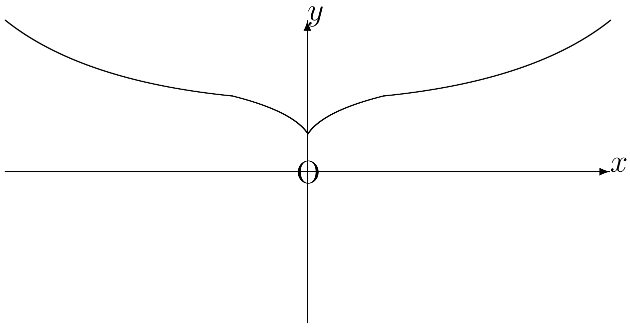
$$f(x) \equiv \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2},$$

in which the first term on the right hand side is unchanged if x is replaced by $-x$ and the second term on the right hand side is reversed in sign if x is replaced by $-x$.

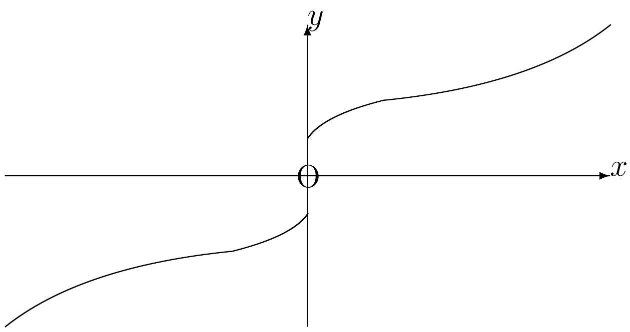
We have thus expressed $f(x)$ as the sum of an even function and an odd function.

(ii) GRAPHS OF EVEN AND ODD FUNCTIONS

(i) The graph of the relationship $y = f(x)$, where $f(x)$ is **even**, will be symmetrical about the y-axis since, for every point (x, y) on the graph, there is also the point $(-x, y)$.



(ii) The graph of the relationship $y = f(x)$, where $f(x)$ is **odd**, will be symmetrical with respect to the origin since, for every point (x, y) on the graph, there is also the point $(-x, -y)$. However, odd functions are more easily recognised by noticing that the part of the graph for $x < 0$ can be obtained from the part for $x > 0$ by reflecting it first in the x-axis and then in the y-axis.

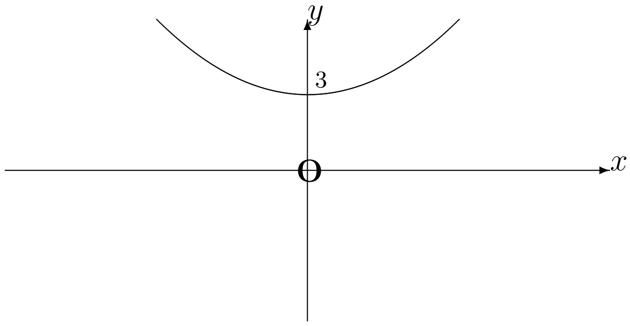


EXAMPLE

Sketch the graph, from $x = -3$ to $x = 3$, of the even function, $f(x)$, defined in the interval $0 < x < 3$ by the formula

$$f(x) \equiv 3 + x^3.$$

Solution



ALGEBRAIC PROPERTIES OF ODD AND EVEN FUNCTIONS

1. The product of an even function and an odd function is an odd function.

Proof:

If $f(x)$ is even and $g(x)$ is odd, then

$$f(-x).g(-x) \equiv f(x).[-g(x)] \equiv -f(x).g(x).$$

2. The product of an even function and an even function is an even function.

Proof:

If $f(x)$ and $g(x)$ are both even functions, then

$$f(-x).g(-x) \equiv f(x).g(x).$$

3. The product of an odd function and an odd function is and even function.

Proof:

If $f(x)$ and $g(x)$ are both odd functions, then

$$f(-x).g(-x) \equiv [-f(x)].[-g(x)] \equiv f(x).g(x).$$

EXAMPLE

Classify the function $f(x) \equiv \sin^4 x. \tan x$ as even, odd or neither even nor odd.

Solution

$$f(-x) \equiv \sin^4(-x). \tan(-x) \equiv \sin^4 x.[-\tan x] \equiv -\sin^4 x. \tan x.$$

The function, $f(x)$, is therefore odd.

10.1.7 EXERCISES

1. If
$$f(x) \equiv 3 + \cos^2\left(\frac{x}{2}\right)$$
determine the values of $f(0), f(\pi), f(\frac{\pi}{2})$.

2. If
$$f(x) \equiv 2x \quad \text{and} \quad g(x) \equiv x^2,$$
verify that
$$f[g(x)] \not\equiv g[f(x)].$$

3. If
$$f(x) \equiv x^2 - 4x + 6,$$
verify that
$$f(2 - x) \equiv f(2 + x).$$

4. Determine simple functions $f(x)$ and $g(x)$ such that the following functions can be identified with $f[g(x)]$:

   (a)
   $$3(x^2 + 2)^3;$$

   (b)
   $$(x^2 + 1)^{-\frac{1}{2}};$$

   (c)
   $$\cos^2 x.$$

5. Evaluate the following limits:

   (a)
$$\lim_{x \to 1} \frac{x^2 + 2x - 3}{x^2 - 8x + 7};$$

   (b)
$$\lim_{x \to 2} \frac{x - 2}{x^2 - (x + 2)};$$

   (c)
$$\lim_{x \to \infty} \frac{3x^2 + 5x - 4}{5x^2 - x + 7};$$

   (d)
$$\lim_{r \to -\infty} \frac{(2r + 1)^2}{(r - 1)(r + 3)};$$

   (e)
$$\lim_{\theta \to 0} \frac{\sin 3\theta}{\theta}.$$

6. Express the function e^x as the sum of an even function and an odd function.

7. Sketch the graph, from $x = -5$ to $x = 5$ of the odd function, $f(x)$, defined in the interval $0 < x < 5$ by the formula
$$f(x) \equiv \cos \frac{\pi x}{10}.$$

8. Classify the function
$$f(x) \equiv \tan^2 x + \text{cosec}^3 x.\cos x$$
as even, odd or neither even nor odd.

10.1.8 ANSWERS TO EXERCISES

1. 4, 3, $\frac{7}{2}$.
2. $2x^2 \not\equiv (2x)^2$.

3. Both are identically equal to $x^2 + 2$.

4. (a)

$$f(x) \equiv 3x^3 \quad \text{and} \quad g(x) \equiv x^2 + 2;$$

(b)

$$f(x) \equiv x^{-\frac{1}{2}} \quad \text{and} \quad g(x) \equiv x^2 + 1;$$

(c)

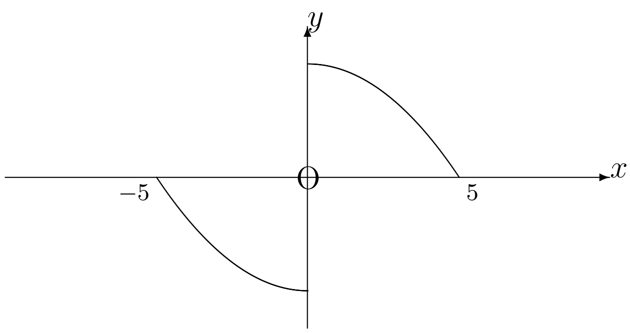
$$f(x) \equiv x^2 \quad \text{and} \quad g(x) \equiv \cos x.$$

5. (a) $-\frac{2}{3}$;
   (b) $\frac{1}{3}$;
   (c) $\frac{3}{5}$;
   (d) 4;
   (e) 3.

6.

$$e^x \equiv \frac{e^x + e^{-x}}{2} + \frac{e^x - e^{-x}}{2}.$$

7. The graph is as follows:



8. The function is neither even nor odd.

"JUST THE MATHS"

UNIT NUMBER

10.2

DIFFERENTIATION 2
(Rates of change)

by

A.J.Hobson

UNIT 10.2 - DIFFERENTIATION 2

RATES OF CHANGE

10.2.1 INTRODUCTION

The functional relationship

$$y = f(x)$$

can be represented diagramatically by drawing the graph of y against x to obtain, in general, some kind of curve.

Between one point of the curve and another, the values of both x and y will change, in general; and the purpose of this section is to introduce the concept of **the rate of increase of y with respect to x**.

A convenient practical illustration which will provide an aid to understanding is to think of y as the distance travelled by a moving object at time x; because, in this case, the rate of increase of y with respect to x becomes the familiar quantity which we know as **speed**.

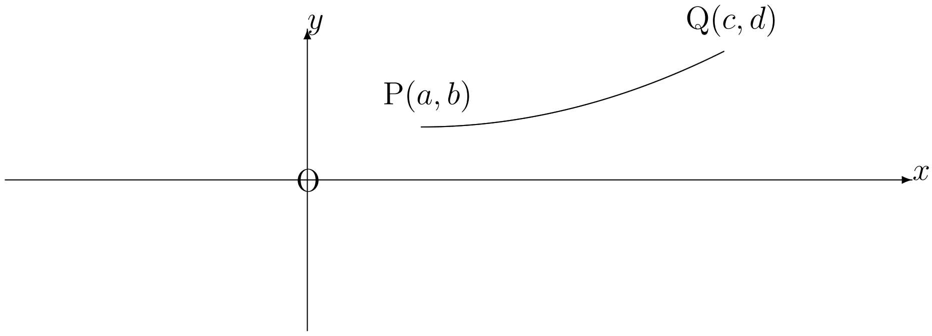
10.2.2 AVERAGE RATES OF CHANGE

Suppose that a vehicle travelled a distance of 280 miles in 7 hours, a journey which is likely to have included short stops, traffic jams, traffic lights and also some fairly high speed motoring. The ratio

$$\frac{280}{7} = 40$$

represents the **"average speed"** of 40 miles per hour over the whole journey. It is a convenient representation of the speed during the journey even though the vehicle might not have been travelling at that speed very often.

Consider now a graph representing the relationship, $y = f(x)$, between two arbitrary variables, x and y, not necessarily time and distance variables.



Between the two points $P(a,b)$ and $Q(c,d)$ an increase of $c-a$ in x gives rise to an increase of $d-b$ in y. Therefore, the average rate of increase of y with respect to x from P to Q is

$$\frac{d-b}{c-a}.$$

If it should happen that y **decreases** as x increases (between P and Q), this quantity will automatically turn out negative; hence,

all rates of increase which are POSITIVE correspond to an INCREASING function,

and

all rates of increase which are NEGATIVE correspond to a DECREASING function.

Note:
For the purposes of later work, the two points P and Q will need to be considered as very close together on the graph, and another way of expressing a rate of increase is to consider notations such as $P(x,y)$ and $Q(x+\delta x, y+\delta y)$ for the pair of points.

Here, we are using the symbols δx and δy to represent **"a small fraction of x"** and **"a small fraction of y"**, respectively. We **do not** mean δ times x and δ times y. We normally consider that δx is positive, but δy may turn out to be negative.

The average rate of increase in this alternative notation is given by

$$\frac{\delta y}{\delta x} = \frac{f(x+\delta x) - f(x)}{\delta x}.$$

In other words,

The average rate of increase is equal to

$$\frac{(\text{new value of } y) \quad \text{minus} \quad (\text{old value of } y)}{(\text{new value of } x) \quad \text{minus} \quad (\text{old value of } x)}$$

EXAMPLE
Determine the average rate of increase of the function

$$y = x^2$$

between the following pairs of points on its graph:

(a) $(3, 9)$ and $(3.3, 10.89)$;

(b) $(3, 9)$ and $(3.2, 10.24)$;

(c) $(3, 9)$ and $(3.1, 9.61)$.

Solution

The results are

(a) $\frac{\delta y}{\delta x} = \frac{1.89}{0.3} = 6.3$;

(b) $\frac{\delta y}{\delta x} = \frac{1.24}{0.2} = 6.2$;

(c) $\frac{\delta y}{\delta x} = \frac{0.61}{0.1} = 6.1$

10.2.3 INSTANTANEOUS RATES OF CHANGE

The results of the example at the end of the previous section seem to suggest that, by letting the second point become increasingly close to the first point along the curve, we could determine the **actual** rate of increase of y with respect to x at the first point only, rather than the **average** rate of increase between the two points.

In the above example, the indications are that the rate of increase of $y = x^2$ with respect to x at the point $(3, 9)$ is equal to 6; and this is called the **"instantaneous rate of increase of y with respect to x"** at the chosen point.

The instantaneous rate of increase in this example has been obtained by guesswork on the strength of just three points approaching $(3, 9)$. In general, we need to consider a limiting process in which an **infinite** number of points approach the chosen one along the curve.

This process is represented by

$$\lim_{\delta x \to 0} \frac{\delta y}{\delta x}$$

and it forms the basis of our main discussion on differential calculus which now follows.

10.2.4 DERIVATIVES

(a) The Definition of a Derivative

In the functional relationship

$$y = f(x)$$

the **"derivative of y with respect to x"** at any point (x, y) on the graph of the function is defined to be the instantaneous rate of increase of y with respect to x at that point.

Assuming that a small increase of δx in x gives rise to a corresponding increase (positive or negative) of δy in y, the derivative will be given by

$$\lim_{\delta x \to 0} \frac{\delta y}{\delta x} = \lim_{\delta x \to 0} \frac{f(x + \delta x) - f(x)}{\delta x}.$$

This limiting value is usually denoted by one of the three symbols

$$\frac{\mathrm{d}y}{\mathrm{d}x}, \quad f'(x) \quad \text{or} \quad \frac{\mathrm{d}}{\mathrm{d}x}[f(x)].$$

Notes:

(i) In the third of these notations, the symbol $\frac{\mathrm{d}}{\mathrm{d}x}$ is called a **"differential operator"**; it cannot exist on its own, but needs to be operating on some function of x. In fact, the first alternative notation is really this differential operator operating on y, which we certainly know to be a function of x.

(ii) The second and third alternative notations are normally used when the derivative of a function of x is being considered without reference to a second variable, y.

(iii) The derivative of a constant function must be zero since the **rate of change** of something which **never changes** is obviously zero.

(iv) Geometrically, the derivative represents the **gradient of the tangent at the point** (x, y) to the curve whose equation is

$$y = f(x).$$

(b) Differentiation from First Principles

Ultimately, the derivatives of **simple** functions may be quoted from a table of standard results; but the establishing of such results requires the use of the definition of a derivative. We illustrate with two examples the process involved:

EXAMPLES

1. Differentiate the function x^4 from first principles.

   Solution

   Here we have a situation where the variable y is not mentioned; so, we could say "let $y = x^4$", and determine $\frac{dy}{dx}$ from first principles in order to answer the question.

   However, we shall choose the alternative notation which does not require the use of y at all.

$$\frac{d}{dx}\left[x^4\right] = \lim_{\delta x \to 0} \frac{(x+\delta x)^4 - x^4}{\delta x}.$$

   Then, from Pascal's Triangle (Unit 2.2),

$$\frac{d}{dx}\left[x^4\right] = \lim_{\delta x \to 0} \frac{x^4 + 4x^3\delta x + 6x^2(\delta x)^2 + 4x(\delta x)^3 + (\delta x)^4 - x^4}{\delta x}$$

$$= \lim_{\delta x \to 0}\left[4x^3 + 6x^2\delta x + 4x(\delta x)^2 + (\delta x)^3\right]$$

$$= 4x^3.$$

   Note:

   This result illustrates a general result which will not be proved here that

$$\frac{d}{dx}[x^n] = nx^{n-1}$$

   for any constant value n, not necessarily an integer.

2. Differentiate the function $\sin x$ from first principles.

   Solution

$$\frac{d}{dx}[\sin x] = \lim_{\delta x \to 0} \frac{\sin(x+\delta x) - \sin x}{\delta x},$$