Basic Mathematics

Contents

1 Basic Skills

This document contains notes on basic mathematics. There are links to the corresponding Leeds University Library *skills@Leeds page*, in which there are subject notes, videos and examples.

If you require more in-depth explanations of these concepts, you can visit the Wolfram Mathworld website:
⇒ **Wolfram link** (*http://mathworld.wolfram.com/*)

- **Algebra (Expanding brackets, Factorising) :**
  ⇒ **Library link**
  (*http://library.leeds.ac.uk/tutorials/maths-solutions/pages/algebra/*).

- **Fractions :**
  ⇒ **Library link**
  (*http://library.leeds.ac.uk/tutorials/maths-solutions/pages/numeracy/fractions.html*).

- **Indices and Powers :**
  ⇒ **Library link**
  (*http://library.leeds.ac.uk/tutorials/maths-solutions/pages/numeracy/indices.html*).

- **Vectors :**
  ⇒ **Library link**
  (*http://library.leeds.ac.uk/tutorials/maths-solutions/pages/mechanics/vectors.html*).

- **Trigonometry and geometry :**
  ⇒ **Library link**
  (*http://library.leeds.ac.uk/tutorials/maths-solutions/pages/trig_geom/*).

- **Differentiation and Integration :**
  ⇒ **Library link**
  (*http://library.leeds.ac.uk/tutorials/maths-solutions/pages/calculus/*).

1.1 Practice Questions

There are practice equations available online to accompany these notes.

2 Linear Algebra

⇒ **Wolfram link** (*http://mathworld.wolfram.com/LinearAlgebra.html*)

2.1 Matrices and Vectors

⇒ **Library link** (*http://library.leeds.ac.uk/tutorials/maths-solutions/pages/mechanics/vectors.html*)

2.1.1 Definitions

A matrix is a rectangular array of numbers enclosed in brackets. These numbers are called entries or elements of the matrix.
e.g.

$$A = \begin{bmatrix} 1 & 3 & 6 \\ 0 & -1 & 4 \end{bmatrix} \tag{1}$$

Matrix A has 2 rows and 3 columns.

A row vector is a matrix with a single row:
e.g.

$$\begin{bmatrix} 1 & 3 & 6 \end{bmatrix} \tag{2}$$

Whereas a column vector is a matrix with a single column:
e.g.

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{3}$$

The size of a matrix is defined by $n \times m$ where n is the number of rows and m is the number of columns. Matrix A, as defined in equation 1, is a 2×3 matrix.

An element of a matrix can be described by its row position and column position. For example: the top left element in matrix A, equal to 1, is in row 1 and column 1 and can be labelled as element a_{11}; the element in the 2^{nd} column of row 1, equal to 3, is labelled as a_{12}. A general element a_{ij} is located in row i and column j (see equation 4 for a further example).

2.1.2 Notation

There are different types of notation for matrices and vectors that you may encounter in text books. Below are some examples:

Matrix	
A	italics
$\boldsymbol{A}$	bold, italics
$\underline{\underline{A}}$	double underline, italics
Vector	
x	italics
$\vec{x}$	top arrow, italics
$\underline{x}$	single underline, italics
$\boldsymbol{x}$	bold

2.1.3 Addition

⇒ **Wolfram link** (*http://mathworld.wolfram.com/MatrixAddition.html*)

⇒ **Video link** (*http://www.youtube.com/watch?v=FX4C-JpTFgY*)

Two matrices (or vectors) of the same size ($n \times m$) may be added together, element by element. For instance, if we have two matrices A and B:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \tag{4}$$

then,

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix} \tag{5}$$

2.1.4 Subtraction

Similar to addition, corresponding elements in A and B are subtracted from each other:

$$A - B = \begin{bmatrix} a_{11} - b_{11} & a_{12} - b_{12} \\ a_{21} - b_{21} & a_{22} - b_{22} \end{bmatrix} \tag{6}$$

2.1.5 Multiplication by a scalar

If λ is a number (i.e. a scalar) and A is a matrix, then λA is also a matrix with entries

$$\begin{bmatrix} \lambda a_{11} & \lambda a_{12} \\ \lambda a_{21} & \lambda a_{22} \end{bmatrix} \tag{7}$$

2.1.6 Multiplication of two matrices

⇒ **Wolfram link** (*http://mathworld.wolfram.com/MatrixMultiplication.html*)

This is non-trivial and is governed by a special rule. Two matrices A , where A is of size $n \times m$, and B of size $p \times q$, can only be multiplied if $m = p$, i.e. the number of columns in A must match the number of rows in B. The matrix produced has size $n \times q$, with each entry being the dot (or scalar) product (see section 2.1.10) of a whole row in A by a whole column in B.

e.g. if

$$A = \begin{bmatrix} 1 & 3 & 6 \\ 0 & -1 & 4 \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{bmatrix} \tag{8}$$

then

$$AB = \begin{bmatrix} (1\times 1) & (1\times 2) & (1\times 3) \\ +(3\times 5) & +(3\times 6) & +(3\times 7) \\ +(6\times 9) & +(6\times 10) & +(6\times 11) \\ \\ (0\times 1) & (0\times 2) & (0\times 3) \\ +(-1\times 5) & +(-1\times 6) & +(-1\times 7) \\ +(4\times 9) & +(4\times 10) & +(4\times 11) \end{bmatrix}$$

$$= \begin{bmatrix} 70 & 80 & 90 \\ 31 & 34 & 37 \end{bmatrix}$$

Formally, if

$$AB = C \qquad \text{then}$$
$$c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj} \tag{9}$$

Aside
When using Matlab (or octave), two matrices can be multiplied in an element-wise sense. This is <u>NOT</u> the same as described above.

2.1.7 Motivation for matrix-matrix multiplication

To understand why we may need to perform matrix-matrix multiplication, consider two customers of a repair garage, Peter and Alex, who require a number of car parts for each of their vehicles. Peter requires 1×3 litre engine and 2 doors, whereas Alex requires 1×5 litre engine and 4 doors. All the parts require a certain number of screws and bolts. But how many total screws and bolts do Peter and Alex need?

We can present the quantity of each car part that Peter and Alex need in a table:

	3 litre engine	5 litre engine	Doors
Peter	1	0	2
Alex	0	1	4

or as the matrix, A:

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 4 \end{bmatrix} . \quad (10)$$

The number of screws and bolts for each car part are expressed in the table:

	bolts	screws
3 litre	3	4
5 litre	1	8
doors	2	1

or can be expressed as matrix, B:

$$A = \begin{bmatrix} 3 & 4 \\ 1 & 8 \\ 2 & 1 \end{bmatrix} . \quad (11)$$

Using simple addition we can find out how many screws and bolts are needed.

1. How many bolts are needed for Peter's car parts?
   $(1 \times 3) + (0 \times 1) + (2 \times 2) = 7.$

2. How many bolts are needed for Alex's car parts?
   $(0 \times 3) + (1 \times 1) + (4 \times 2) = 9.$

3. How many screws are needed for Peter's car parts?
   $(1 \times 4) + (0 \times 8) + (2 \times 4) = 6.$

Or we can use matrix multiplication to get all four scenarios:

$$AB = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 1 & 8 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 7 & 6 \\ 9 & 12 \end{bmatrix}$$

2.1.8 Matrix-vector multiplication

Since a vector is a special case of a matrix, this is simply a special case of the matrix-matrix multiplication we have already discussed. Consider multiplying a column vector of length m by a matrix of size $n \times m$,
e.g.

$$\begin{bmatrix} 1 & 3 & 6 \\ 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 23 \\ 11 \end{bmatrix}$$

which results in a column vector of length n and in this case $n = 2$.

2.1.9 Special Matrices

Identity Matrix, I: The identity matrix, I, of size $n \times n$, is defined in equation 12.

$$a_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{12}$$

i.e. if $n = 2$

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{13}$$

This is a special case of a diagonal matrix possessing non-zero entries only on its diagonal e.g.

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

If A is a square $n \times n$ matrix, then the identity matrix $I_{n\times n}$ has the special property that:

$$AI = IA = A \tag{14}$$

NB: I is the equivalent of 1 in scalar arithmetic i.e. $1 \times 2 = 2 \times 1 = 2$.

Transpose, A^T: If A is a $n \times m$ matrix then the transpose of A, denoted A^T, is a $m \times n$ matrix found by swapping rows and columns of A,
e.g.

$$\text{if} \qquad A = \begin{bmatrix} 1 & 3 & 6 \\ 0 & -1 & 4 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & 0 \\ 3 & -1 \\ 6 & 4 \end{bmatrix}$$

Inverse matrix, A^{-1} If A is an $n \times n$ matrix, sometimes (see later) there exists another matrix called the inverse of A, written A^{-1}, such that

$$AA^{-1} = A^{-1}A = I \tag{15}$$

NB: For scalar numbers, x^{-1} is the inverse of x when considering multiplication, since

$$xx^{-1} = x^{-1}x = 1 \tag{16}$$

Clearly when $x = 0$ this breaks down and x has no inverse — this is also true when dealing with some matrices.

2.1.10 Scalar products and orthogonality

The scalar product (or dot product, or inner product) of two column vectors of length n, where $\underline{x} = (x_1, x_2, x_3, \ldots x_n)$ and $\underline{y} = (y_1, y_2, y_3, \ldots, y_n)$, is

$$\underline{x} \cdot \underline{y} = \sum_{i=1}^{n} x_i \, y_i.$$

This can also be written as $\underline{x}^T \underline{y}$; that is, the product of a row vector of length n with a column vector of length n. Two vectors are said to be orthogonal if their scalar product is zero.

2.2 Linear Systems

⇒ **Wolfram link** (*http://mathworld.wolfram.com/LinearSystemofEquations.html*)

⇒ **Video link** (*http://www.youtube.com/watch?v=ZK3O402wf1c*)

A linear system of equations such as

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

can be written as

$$\begin{bmatrix} 5 & 3 & 1 \\ 2 & -10 & -3 \\ 0 & 4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 7 \end{bmatrix}$$

as can be verified by multiplying out the left hand side. When solving the linear system $A\underline{x} = \underline{b}$, (where A is a matrix, $\underline{x}$ is the vector $(x, y, z, \dots)$ and $\underline{b}$ is a vector of numbers) two cases can arise:

i) A^{-1} exists.

Then $A^{-1}A\underline{x} = A^{-1}\underline{b} \quad \Rightarrow \quad I\underline{x} = A^{-1}\underline{b}$

so $\underline{x} = A^{-1}\underline{b}$

ii) A^{-1} doesn't exist. There is then no solution in general.

2.3 Determinants

⇒ **Wolfram link** (*http://mathworld.wolfram.com/Determinant.html*)

⇒ **Video link** (*http://www.youtube.com/watch?v=23LLB9mNJvc*)

How do we know when A^{-1} exists? One method is to calculate the determinant of A, written det A or $\mid A \mid$. The determinant is a single number that contains enough information about A to determine whether it is invertible.

2×2 determinants: In the 2×2 case, if

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

then $\mid A \mid = ad - bc$. For example, for the linear system given by

$$x + 2y = 2$$
$$3x + 4y = 3$$

the determinant of the coefficient matrix is

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 4 - 6 = -2 \neq 0 \tag{18}$$

The matrix A is therefore invertible (see section 2.3.1) and so the solution exists.
As another example, consider

$$x + 2y = 2$$
$$2x + 4y = 3$$

The determinant of the coefficient matrix is

$$\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 4 - 4 = 0 \tag{19}$$

Therefore, A has no inverse and so no solution exists. This can also be seen in the fact that it is not possible that $x + 2y$ simultaneously equal both 2 and 3.

3×3 determinants: Determinants can be generalised to $n \times n$ matrices. For the 3×3 matrix A,

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

the determinant of A is:

$$\mid A \mid = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

or equivalently,

$$\mid A \mid = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

e.g.

$$\begin{vmatrix} 1 & 0 & 0 \\ 2 & 3 & 4 \\ 5 & 6 & 7 \end{vmatrix} = 1 \begin{vmatrix} 3 & 4 \\ 6 & 7 \end{vmatrix} = 21 - 24 = -3$$

or

$$\begin{vmatrix} 1 & 0 & 0 \\ 2 & 3 & 4 \\ 5 & 6 & 7 \end{vmatrix} = 1 \begin{vmatrix} 3 & 4 \\ 6 & 7 \end{vmatrix} - 2 \begin{vmatrix} 0 & 0 \\ 6 & 7 \end{vmatrix} + 5 \begin{vmatrix} 0 & 0 \\ 3 & 4 \end{vmatrix} = -3$$

Do whichever is easier!

2.3.1 Using determinants to invert a 2×2 matrix

The determinant can be used in finding the inverse of a 2×2 matrix.

For $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the inverse can be found using the formula

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{20}$$

For example: find the inverse of matrix $B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

$$B^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^{-1} = \frac{1}{-2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}.$$

2.4 Eigenvalues and Eigenvectors

⇒ **Wolfram link 1** (*http://mathworld.wolfram.com/Eigenvalue.html*)

⇒ **Wolfram link 2** (*http://mathworld.wolfram.com/Eigenvector.html*)

⇒ **Video link** (*http://www.youtube.com/watch?v=lXNXrLcoerU*)

Often we are interested in whether a matrix can stretch a vector. In such a case:

$$Av = \lambda v \ , \tag{21}$$

where λ is the "stretch factor". The scalar λ is called an eigenvalue (from the German: eigen meaning same) and v is an eigenvector. $Av = \lambda v$ is equivalent to:

$$(A - \lambda I)v = 0 \ . \tag{22}$$

If $det(A - \lambda I) \neq 0$ then the system can be solved to find $v = (A - \lambda I)^{-1}0 = 0$. If we want non-zero vectors v, then we require $|A - \lambda I| = 0$.

To find the eigenvectors and eigenvalues, we use a two stage process:

i) Solve $|A - \lambda I| = 0$,

ii) Find v.

For example:

$$A = \begin{bmatrix} 14 & -10 \\ 5 & -1 \end{bmatrix}$$

i) The eigenvalues λ are such that

$$\left| \begin{pmatrix} 14 & -10 \\ 5 & -1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| = 0$$

$$\text{or} \qquad \begin{vmatrix} 14-\lambda & -10 \\ 5 & -1-\lambda \end{vmatrix} = 0$$

or $(14-\lambda)(-1-\lambda)-(-10)(5)=0$

$\Rightarrow \lambda^2 - 13\lambda + 36 = 0$

$\Rightarrow (\lambda-4)(\lambda-9)=0$ so $\lambda = 4$ or $\lambda = 9$

ii) Now to find the eigenvectors, v:

If $\lambda = 4$ then

$$(A-4I)\begin{pmatrix} x \\ y \end{pmatrix} = 0 \qquad \text{or}$$

$$\begin{pmatrix} 10 & -10 \\ 5 & -5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0 \qquad \Rightarrow \qquad \text{x=y}$$

Hence the eigenvector is: $\binom{x}{x} = x\binom{1}{1}$, or $y\binom{1}{1}$, or just $\binom{1}{1}$, (see below).

If $\lambda = 9$, then

$$\begin{pmatrix} 5 & -10 \\ 5 & -10 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{so} \quad x = 2y$$

giving the eigenvector of $\binom{2y}{y}$ or just $\binom{2}{1}$.

The eigenvectors provide a direction and so we can ignore common factors. This is because if v is an eigenvector, then so is μv for any value of μ.
i.e. $Av = \lambda v \Rightarrow A(\mu v) = \lambda(\mu v)$.

For example: if $\lambda = 4$ above, then $\binom{1}{1}, \binom{2}{2}, \binom{4}{4}$ are <u>all</u> eigenvectors. We typically choose the simplest!

Matrix diagonalisation Suppose A is a $n \times n$ matrix. Form a new matrix V, each column of which is an eigenvector of A. If V^{-1} exists then,

$V^{-1}AV = \Lambda$ or $A = V\Lambda V^{-1}$ where Λ is a diagonal matrix with $(\lambda_1, \lambda_2, \dots)$, the eigenvalues of A on the diagonal.

E.g.

$$A = \begin{pmatrix} 14 & -10 \\ 5 & -1 \end{pmatrix} \quad \text{then} \quad V = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix}$$

NB - Eigenvector 1 is: $\binom{1}{1}$, eigenvector 2 is: $\binom{2}{1}$

It turns out that V^{-1} exists and is:

$$V^{-1} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$$

$$\left[\text{Check:} \quad VV^{-1} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]$$

So then the matrix product

$$\begin{aligned} V^{-1}AV &= \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 14 & -10 \\ 5 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 4 & 18 \\ 4 & 9 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix} \end{aligned}$$

is equal to the matrix Λ.

Special case: Symmetric matrix A matrix A is symmetric if $A = A^T$. It turns out that, in this case, all the eigenvectors are orthogonal, i.e. if V_1 and V_2 are different eigenvectors, then $V_1 \cdot V_2 = 0$.
If each eigenvector is normalised such that $V_i \cdot V_i = 1$, then,
$V^T = V^{-1}$ (This is an easy way to find V^{-1}).

Example:

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \qquad (A^T = A)$$

Eigenvectors λ are given by:

$$(1-\lambda)^2 - 4 = 0$$
$$\lambda^2 - 2\lambda - 3 = 0 \text{ or } (\lambda - 3)(\lambda + 1) = 0$$
$$\text{so,} \qquad \lambda = 3 \text{ or } \lambda = -1$$

$\lambda = -1$:

$$\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \Rightarrow \qquad x + y = 0$$

Hence, eigenvector is: $\binom{1}{-1}$.

But this can be normalised (since it's magnitude is arbitrary) by any number μ : $\mu\binom{1}{-1}$.

Let's choose μ such that: $v^T v = 1,$ where $v = \mu\binom{1}{-1}$.

$$\Rightarrow \qquad 1 = \mu^2 \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 2\mu^2.$$

Hence, $\mu^2 = \frac{1}{2}$ *or* $\mu = \frac{1}{\sqrt{2}}$
Our normalised eigenvector is then $\frac{1}{\sqrt{2}}\binom{1}{-1}$.

$\underline{\lambda = 3}$:

$$\begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad -2x + 2y = 0$$

or, $x = y$

Hence, eigenvector is $\mu\binom{1}{1}$.

Normalise such that $v^T v = 1$, so

$$\mu^2 \begin{pmatrix} 1 & 1 \end{pmatrix} \binom{1}{1} = 1 \quad \Rightarrow \quad \mu = \frac{1}{\sqrt{2}} \quad \text{as before.}$$

The normalised eigenvector is then $\frac{1}{\sqrt{2}}\binom{1}{1}$.

Define:

$$V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \qquad \Lambda = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$$

Then we can check that

$$
\begin{aligned}
V\Lambda V^{-1} \quad &= \quad \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\
&= \quad \frac{1}{2}\begin{pmatrix} -1 & 3 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\
&= \quad \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \quad = \quad A
\end{aligned}
$$

Why is this useful?

<u>Example:</u> What is A^8 ? Using the matrix diagonalisation,

$$
\begin{aligned}
A^8 &= \left(V\Lambda V^{-1}\right)\left(V\Lambda V^{-1}\right)\left(V\Lambda V^{-1}\right)\left(V\Lambda V^{-1}\right)\left(V\Lambda V^{-1}\right)\left(V\Lambda V^{-1}\right)\left(V\Lambda V^{-1}\right)\left(V\Lambda V^{-1}\right) \\
&= V\Lambda\left(V^{-1}V\right)\Lambda\left(V^{-1}V\right)\Lambda\left(V^{-1}V\right)\Lambda\left(V^{-1}V\right)\Lambda\left(V^{-1}V\right)\Lambda\left(V^{-1}V\right)\Lambda\left(V^{-1}V\right)\Lambda\left(V^{-1}V\right)V^{-1} \\
&= \left(V\,\Lambda^8\,V^{-1}\right).
\end{aligned}
$$

and Λ^8 is an easy matrix to raise to a power:

$$
\begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}^8 = \begin{pmatrix} 1 & 0 \\ 0 & 3^8 \end{pmatrix} \tag{23}
$$

This method is much easier than multiplying the matrix by itself 8 times.

3 Differentiation and Integration

⇒ **Library link** (*http://library.leeds.ac.uk/tutorials/maths-solutions/pages/calculus/*)

3.1 Differentiation

⇒ **Wolfram link** (*http://mathworld.wolfram.com/Derivative.html*)

Suppose the function $f(t)$ gives distance as a function of time as shown in Figure 1.

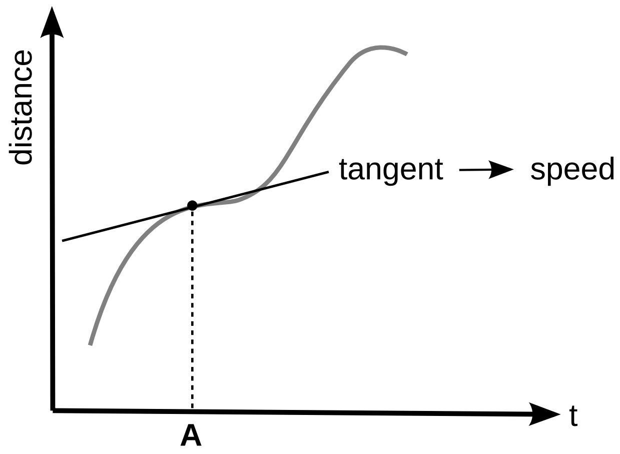

Figure 1

At point A, the sketched tangent gives the instantaneous rate of change of distance with time or speed.

To calculate the speed (or gradient function) at any time, we approximate the tangent by connecting two neighbouring points: $f(t)$ and $f(t+h)$ (see Figure 2).

An estimate of the gradient at time t is then

$$\frac{f(t+h)-f(t)}{(t+h)-t} = \frac{f(t+h)-f(t)}{h} \tag{24}$$

As $h \to 0$, this estimate becomes more accurate.

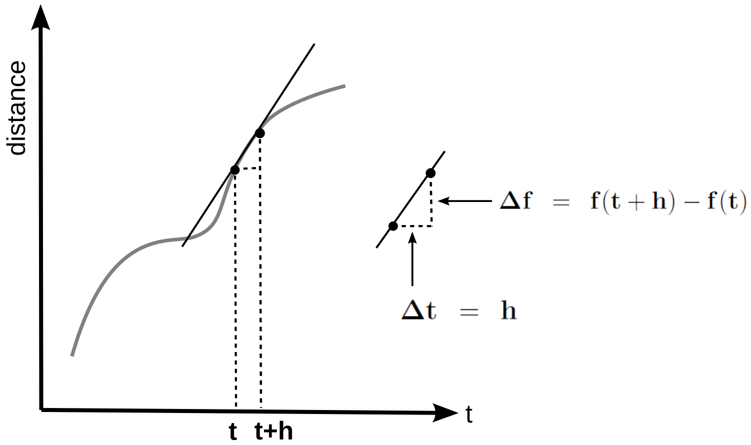

Figure 2

For example   Suppose $f(t) = t^2$, then

$$\begin{aligned}\frac{f(t+h)-f(t)}{h} &= \frac{(t+h)^2-t^2}{h} \\ &= \frac{2th+h^2}{h} \\ &= 2t+h\end{aligned}$$

therefore: $$\frac{f(t+h)-f(t)}{h} \to 2t \quad \text{as} \quad h \to 0$$

Hence the gradient of f at any point t is $2t$.

3.1.1 Notation

The gradient or derivative of a function $f(t)$ can be written:

$$\frac{df}{dt}, \quad f' \quad or \quad \dot{f} \tag{25}$$

We can also have higher derivatives. Consider the gradient of a gradient function. If f represents distance, then f' is the speed and f'' is the acceleration.

3.1.2 Standard Results

f	f'
1	0
t	1
t^n	nt^{n-1}
t^{-m}	$-mt^{-(m+1)}$
$\sin t$	$\cos t$
$\cos t$	$-\sin t$
e^t	e^t
$\ln t$	$\frac{1}{t}$

Differentiation is linear: e.g.

$$f(t) = 3\ln t + 9\sin t$$
$$f'(t) = \frac{3}{t} + 9\cos t$$

3.1.3 Product rule

If we need to take the time derivative of the product uv of two functions $u(t)$ and $v(t)$ then we use the product rule.

$$\frac{d}{dt}(uv) = u'v + v'u$$

3.1.4 Chain rule

The chain rule can be used to differentiate more complicated functions.
The chain rule is defined as:

$$\frac{df}{dt} = \frac{df}{du}\frac{du}{dt}$$

For example:

1. How do we differentiate the function $f(t) = \sin 2t$?
   We know how to differentiate $f = \sin u$, so let's define $u = 2t$. Then we simply need to assemble the ingredients for the chain rule: $\frac{du}{dt} = 2$ and $\frac{df}{du} = \cos u$. It then follows that

$$\begin{aligned} \frac{df}{dt} &= \frac{df}{du}\frac{du}{dt} \\ &= 2\cos(u) \\ &= 2\cos(2t) \end{aligned}$$

2. How do we differentiate the function $f(t) = \frac{1}{u(t)}$?
   We know how to differentiate $f = 1/u$, so let's go ahead and use the chain rule, the key ingredient we need being $\frac{df}{du} = -u^{-2} = -\frac{1}{u^2}$. Then

$$\begin{aligned} \frac{df}{dt} &= \frac{df}{du}\frac{du}{dt} \\ &= -\frac{1}{u^2}\frac{d(u(t))}{dt} \\ &= -\frac{u'}{u(t)^2} \end{aligned}$$

3.1.5 Quotient rule

If we need to take the derivative of a quotient of functions: u/v, then we use the quotient rule.

$$\begin{aligned} \frac{d}{dt}\left(\frac{u}{v}\right) &= \frac{d}{dt}\left(uv^{-1}\right) \\ &= \frac{du}{dt}v^{-1} - \frac{dv}{dt}\frac{u}{v^2} \\ &= \frac{vu' - uv'}{v^2} \end{aligned}$$

3.1.6 Stationary points in 1D

⇒ **Wolfram link** (*http://mathworld.wolfram.com/StationaryPoint.html*)

For a function $f(x)$, a stationary point is where the gradient, $f'(x)$ vanishes. To decide whether it represents a local maximum or minimum, we need the 2nd derivative. If $f''(x) > 0$ at the stationary point, then it is a minimum; conversely if $f''(x) < 0$ then it's a maximum.

For example, find and classify the stationary points of $f(x) = \sin x$ on $[0, 2\pi]$. The gradient function is $f'(x) = \cos x$, which on the range in question is zero at $x = \pi/2$ and $x = 3\pi/2$. Evaluating the 2nd derivative, $f''(x) = -\sin x$ at these points gives $f''(\pi/2) = -1 < 0$ and $f''(3\pi/2) = 1 > 0$. Hence $x = \pi/2$ is a local maximum, and $x = 3\pi/2$ is a local minimum.

3.1.7 Partial derivatives

⇒ **Wolfram link** (*http://mathworld.wolfram.com/PartialDerivative.html*)

Suppose f is a function of more than one independent variable, e.g. $f = xy$.
We can make sense of the gradient by varying each variable, one at a time:

$$\frac{\partial f}{\partial x} = y \qquad\qquad \frac{\partial f}{\partial y} = x$$

hold y constant hold x constant

The notation $\frac{\partial f}{\partial v}$ means a "partial" derivative with respect to v and regards all other variables as constant.
e.g. $f = x + x\sin y;$ $\frac{\partial f}{\partial x} = 1 + \sin y,$ $\frac{\partial f}{\partial y} = x\cos y.$

We can take 2$^{\text{nd}}$ and higher order derivatives:
e.g.

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\frac{\partial f}{\partial x} \\ \frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x}\frac{\partial f}{\partial y} \qquad \left(= \frac{\partial}{\partial y}\frac{\partial f}{\partial x}\right) \end{aligned} \tag{26}$$

3.1.8 Stationary points in 2 dimensions

One mathematical description of a surface in 3D is a function f, giving the height as a function of coordinates x and y. Just as in 1D, stationary points are where the gradient in all directions is zero (corresponding to local maxima, minima or saddle points) and are given by the conditions

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0.$$

3.1.9 Taylor Series

We can approximate the behaviour of the function at a point using knowledge of its derivatives. For example, suppose a car is at a distance of 50 km and is travelling at 100 km/hr. Where will it be in 30 minutes?

If its speed is constant, $50 + \frac{100}{2} = 100$ km. But if its speed changes, then we need a correction term.

A function $f(a+h)$ is related to its behaviour at $x = a$ by,

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2!}f''(a) + \frac{h^3}{3!}f'''(a) + ... \tag{27}$$

For example: If $f(x) = \sin(x)$ and $a = 0$,

$$\begin{aligned}\sin(0+h) &= \sin(0) + h\sin'(0) + \frac{h^2}{2!}\sin''(0) + \frac{h^3}{3!}\sin'''(0) + ...\\ &= h - \frac{h^3}{3!} + ...\end{aligned}$$

We can generalise this idea to functions of several variables. If $f(x,y)$ is expanded near (a,b):

$$f(a+h, b+k) = f(a,b) + h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y} + \text{higher order terms}$$

e.g. If $f(x,y) = \sin(xy)$, how does f behave close to $(1,\pi)$?

$$
\begin{aligned}
f(1,\pi) &= 0; \\
\frac{\partial f}{\partial x} = y\cos(xy)\bigg|_{(1,\pi)} &= -\pi \\
\frac{\partial f}{\partial y}\bigg|_{(1,\pi)} &= -1
\end{aligned}
$$

Hence, $f(1+h,\pi+k) = -\pi h - k$, to 1st order.

3.2 Integration

⇒ **<u>Wolfram link</u>** (*http://mathworld.wolfram.com/Integral.html*)

The integral of a function f between $x = a$ and $x = b$ is the area under the curve of f: $\int_a^b f(x)dx$. To find it, we often use the "fundamental theorem" of calculus:

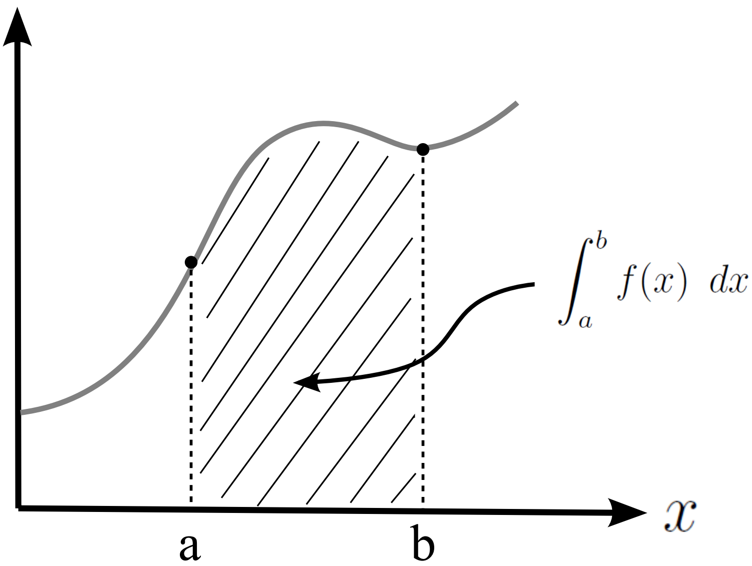

Figure 3

$$\int_a^b f(x)\ dx \quad = \quad \big[F(x)\big]_a^b \quad = \quad F(b) - F(a) \tag{28}$$

where $F'(x) = f(x)$

That is, F is the anti-derivative of f.

$$\text{e.g.} \int_1^2 x^2\ dx \quad = \quad \left[\frac{1}{3}x^3\right]_1^2 \quad = \quad \frac{1}{3}(8-1) \quad = \quad \frac{7}{3}$$

$$\text{since, } \frac{d}{dx}\left(\frac{1}{3}x^3\right) \;=\; x^2.$$

An indefinite integral is an integral without limits and gives a function that is the anti-derivative of f (including an arbitrary constant):

$$\text{e.g.} \int x^2\ dx \quad = \quad \frac{1}{3}x^3 + C$$

Standard Results

$f(x)$	$\int f(x)dx$
1	$x + C$
x^n	$\frac{1}{n+1}x^{n+1} + C$
$\cos x$	$\sin x + C$
$\sin x$	$-\cos x + C$
e^x	$e^x + C$
$\frac{1}{x}$	$\ln \lvert x \rvert + C$

3.2.1 Finding Integrals

1. By parts

   We have already seen that: $(uv)' = u'v + v'u$.

   If we integrate this:

$$\begin{aligned} \int (uv)'dx &= \int u'v + v'u \\ \Rightarrow uv &= \int u'v + v'u \\ \text{so,} \int u'v &= uv - \int v'u \end{aligned} \tag{29}$$

   e.g. Find:

$$\int x \sin(x) \ dx$$

   Let:

$$f = \int x \sin(x) \ dx, \qquad v = x, \qquad \text{and} \qquad u' = \sin x.$$

Therefore:

$$
\begin{aligned}
f &= -\cos x x - \int 1(-\cos x) \ dx \\
&= -\cos x x + \sin x + C
\end{aligned}
$$

$\Big[$Check! $\quad \frac{d}{dx}(-x\cos x + \sin x) \quad = \quad -\cos x + x\sin x + \cos x \quad = \quad x\sin x\Big]$

2. By substitution

For example, suppose we want to find

$$\int x\cos\left(x^2+1\right) \ dx$$

Write

$$u = x^2 + 1, \tag{30}$$

so that $\frac{du}{dx} = 2x$ or rearranging to give $dx = \frac{du}{2x}$. Substituting all 'x' variables for 'u' variables gives

$$
\begin{aligned}
\int x\cos(x^2+1) \ dx &= \int x\cos u \ \frac{du}{2x} \\
&= \int \frac{1}{2}\cos u \ du \\
&= \frac{1}{2}\sin u + C \\
&= \frac{1}{2}\sin\left(x^2+1\right) + C.
\end{aligned}
$$

As another example, suppose we want to find:

$$\int_0^1 (1-x^2)^{-\frac{1}{2}} \ dx \tag{31}$$

Choose $x = \sin u$, so that $dx = \cos u \, du$. The integral then becomes,

$$\int_0^1 \frac{1}{\cos u} \cos u \;\; du \quad = \quad \big[u\big]_0^1 \quad = \quad 1.$$

4 Complex Numbers

⇒ **Library link** (*http://library.leeds.ac.uk/tutorials/maths-solutions/pages/complex_numbers/*)

⇒ **Wolfram link** (*http://mathworld.wolfram.com/ComplexNumber.html*)

⇒ **Video link** (*http://ocw.mit.edu/resources/res-18-008-calculus-revisited-complex-variables-differential-equations-and-linear-algebra-fall-2011/part-i/lecture-1-the-complex-numbers/*)

⇒ **Video link** (*http://archive.org/details/mit-ocw-18.03-lec06*)

4.1 Motivation

It was not so long ago that equations like $x + 5 = 0$ or $2x = 3$ were considered absurd - how can you have -5 or $\frac{3}{2}$ cows? Nevertheless, these mathematical representations of negative and fractional numbers are extremely useful.

Along similar lines, it wasn't very long ago that the equation: $x^2 = -1$, was considered ridiculous, for how can a squared number be negative? But if we proceed anyway and define a solution, it turns out to be very useful indeed.

4.1.1 Graphical concept

How might we visualise the solution to: $x^2 \times 1 = 9$? That is, what *transformation* applied twice to 1, gives 9? The answer of course is 3 or -3.

But let's consider, $x^2 \times 1 = -1$. What transformation, applied twice to 1, gives -1?

We can express this as a rotation of $2 \times 90°$, (see Figure 4).

A rotation of $2 \times -90°$ also works as in Figure 5,

To make sense of this, we need to define a new "imaginary dimension"; i or $-i$ is what becomes of 1 or -1 after a rotation of $90°$ or $-90°$.

What happens if we keep multiplying by i?

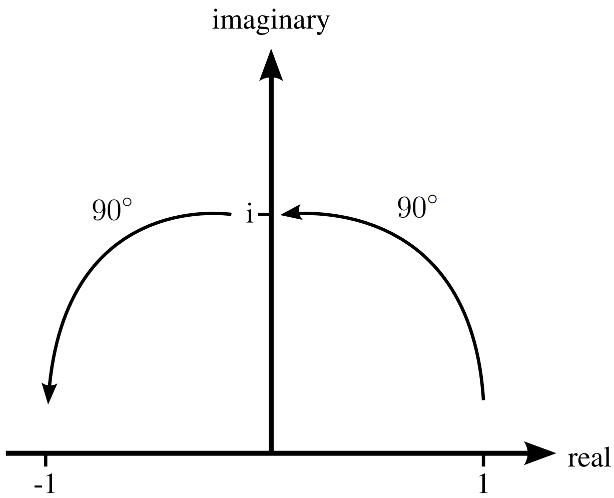

Figure 4

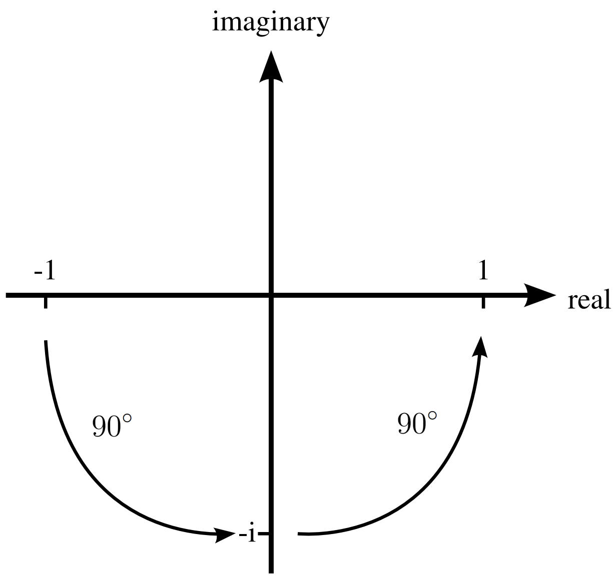

Figure 5

$1, i, i^2, i^3, i^4, \cdots = 1, i, -1, -i, 1, i, -1, -i, 1, \ldots$ as depicted below in Figure 6.

4.2 Definition

A **complex number** z is written (x, y) or $x + iy$ where x is the real part of z and y is the imaginary part. The number i satisfies $i^2 + 1 = 0$. For any complex number z, its real part x is $x = \Re(z)$ or $Re(z)$, and the complex (or imaginary) part y, is $y = \Im(z)$ or $Im(y)$. If $y = 0$,

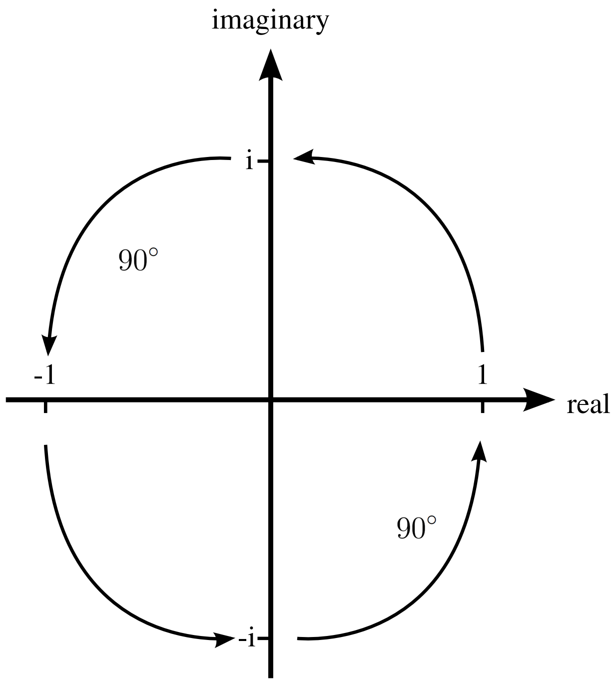

Figure 6

then z is purely real; if $x = 0$, then z is purely imaginary.

4.3 Complex Plane

We can plot a complex number in an $x - y$ domain called the *complex plane* or the *Argand Diagram*. For example, $z = x + iy$ is displayed in the Argand diagram in Figure 7.

4.4 Addition/Subtraction

Two complex numbers $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$ can be added or subtracted by adding or subtracting their real and imaginary parts separately.
For example:

$$(5 + 2i) + (7 + 9i) = 12 + 11i$$

$$(2 + 2i) - (3 + 4i) = -1 - 2i$$

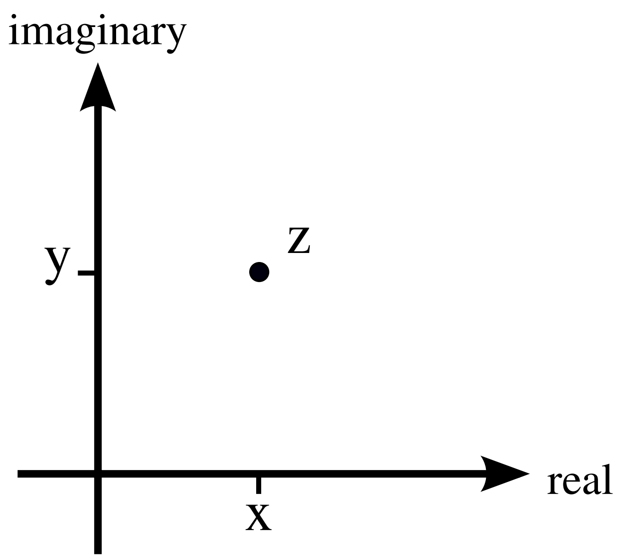

Figure 7: An argand diagram showing the complex plane. The complex number z has real part x and complex part y.

4.5 Multiplication

For two complex numbers z_1 and z_2, find their product by multiplying out in full.

For example:

$$\begin{aligned}(5+2i)\,(1+3i) &= 5+6i^2+17i \\ &= -1+17i,\end{aligned}$$

using $i^2 = -1$.

4.6 Conjugates

If z is a complex number $x + iy$, then the conjugate to z is: $\bar{z} = x - iy$.

This is useful because:

$$\begin{aligned}z\bar{z} &= (x+iy)(x-iy) \\ &= x^2+(-i^2)y^2 \\ &= x^2+y^2, \end{aligned} \tag{32}$$

is a purely real number.

Geometrically, $z\bar{z} = |z|^2$, where $|z|$ is the "magnitude" of z, see Figure 8 below.

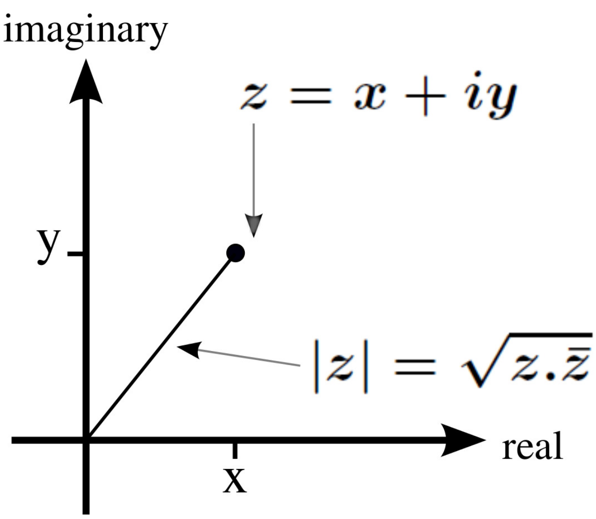

Figure 8: An argand diagram showing the magnitude $|z| = \sqrt{z\bar{z}}$ of the complex number $z = x + iy$.

4.7 Division

It is not immediately obvious how to divide two complex numbers z_1 and z_2. However, we do know how to divide a complex number by a real number. For example:

$$\frac{1}{2}(2 + 2i) = 1 + i.$$

To divide z_1 by z_2 we need to use $\bar{z}_2$:

$$\frac{z_1}{z_2} = \frac{z_1 \bar{z}_2}{z_2 \bar{z}_2} = \frac{(z_1 \bar{z}_2)}{|z_2|^2} \tag{33}$$

For example:

$$\begin{aligned} \frac{1+i}{2+3i} &= \frac{(1+i)(2-3i)}{(2+3i)(2-3i)} \\ &= \frac{5-i}{13} \\ &= \frac{5}{13} - \frac{1}{13}i \end{aligned}$$

4.8 Polar Form

A useful representation of a complex number is in polar coordinates.

For example: the complex number: $2+i$ as shown in Figure 9 can also be represented by (r, θ) , where r is the magnitude of z (the distance from 0) and θ is the angle with the horizontal.

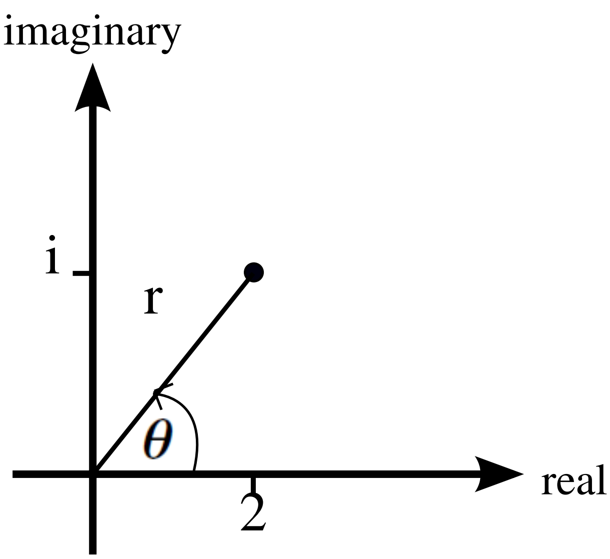

Figure 9: An argand diagram showing the magnitude (or modulus) r and angle (or argument) θ of a complex number.

In this notation,

$x = r\cos\theta$ and $y = r\sin\theta$

so, $x + iy = r(\cos\theta + i\sin\theta)$.

The magnitude of z is also called the **modulus**, $|z|$; θ is called the **argument** and denoted:

$$\theta = \arg(z) = \tan^{-1}\left(\frac{y}{x}\right). \tag{34}$$

In order that θ is unique, $-\pi \leq \arg(z) \leq \pi$ is often used.

4.9 Exponential Notation

It turns out that:

$$z = r(\cos\theta + i\sin\theta) \qquad \text{can also be written as}$$

$$z = re^{i\theta}, \tag{35}$$

$$\tag{36}$$

$$\text{where} \qquad e^{i\theta} = \cos\theta + i\sin\theta \tag{37}$$

This makes it easy to multiply and divide in polar form:

$$\begin{aligned} z_1 z_2 &= r_1 e^{i\theta_1} r_2 e^{i\theta_2} &= r_1 r_2 e^{i(\theta_1+\theta_2)}, \\ \frac{z_1}{z_2} &= r_1 e^{i\theta_1} r_2^{-1} e^{i\theta_2} &= \frac{r_1}{r_2} e^{i(\theta_1-\theta_2)}, \end{aligned}$$

and compute powers, e.g.

$$\begin{aligned} z_1^3 &= \left(r_1 e^{i\theta_1}\right)^3, \\ &= r_1^3 e^{3i\theta_1}. \end{aligned}$$

4.10 Application to waves

Some links about waves as a refresher:

⇒ **Movie link** Sine and cosine waves
(*http://videos.kightleys.com/Science/Maths/23131008_CsD3fs/1880848370_VMGSWd3#!i=1880848370&k=VMGSWd3*)

⇒ **Movie link** Superposition of waves
(*http://www.acs.psu.edu/drussell/demos/superposition/superposition.html*)

Waves can be described (in real form) as

$$y(t) = A\cos(\omega t + \phi) \tag{38}$$

where A is the amplitude, ω is the angular frequency and ϕ is the phase.

However, it has been established that there is a close relationship between complex numbers and sine/cosine functions — it can be useful to use complex numbers to express wave characteristics. Since waves are a physical phenomenon, we would like any representation to be real. We can therefore speak of "complex waves", subject to the expectation that we are only really interested in the real part (we often just throw away the imaginary part). Therefore, the complex wave

$$y = Ae^{i(\omega t+\phi)} = A\cos(\omega t + \phi) + iA\sin(\omega t + \phi) \tag{39}$$

has real part $A\cos(\omega t + \phi)$ and so the two representations: (38) and (39) are essentially equivalent.

If a wave is described by $y = Ce^{it}$ where C is a complex number, then in fact C contains information about both amplitude and phase. This is because we can write $C = Ae^{i\phi}$, so that

$$\begin{aligned} y &= Ae^{i\phi}e^{it} \\ &= Ae^{i(t+\phi)} \\ &= A\cos(t+\phi) + iA\sin(t+\phi) \end{aligned} \tag{40}$$

The amplitude of the wave is A and the phase is ϕ.

Examples of two types of waves are shown in Figure 10. Both waves have amplitude of 2 but the phase is shifted. Visually you can see that one wave ($-\sin t$ wave shown as dashed lines) can be shifted by $\pi/2$ to be equal to the $\cos t$ wave.

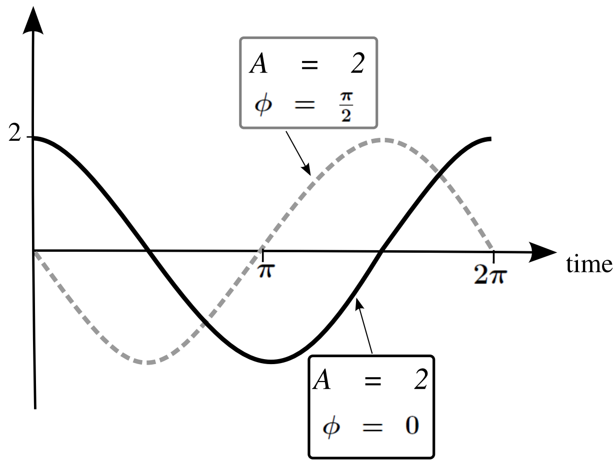

Figure 10

Aside:
Both θ and ϕ have been used as angles in this chapter on complex numbers and the general convention is to use:

- θ for angles (in the polar plane);
- ϕ for phase of waves.

4.10.1 Amplitude and phase

Links:
⇒ **Wiki link** Phase (*http://en.wikipedia.org/wiki/Phase_(waves)*)
⇒ **Wolfram link** Amplitude (*http://mathworld.wolfram.com/Amplitude.html*)

To find the amplitude and phase of a wave function we convert the function into the exponential complex form.
For example:

- $\cos t$ is the real part of $e^{(it+0)} = e^{it}e^0$ and has amplitude 1 and phase 0.

- $2\sin t$ is the real part of $2e^{i(t-\frac{\pi}{2})} = 2e^{it}e^{-i\pi/2}$ and has amplitude 2 and phase $-\pi/2$.
- $-\cos t$ is the real part of $e^{i(t+\pi)} = e^{it}e^{i\pi}$ and has amplitude 1 and phase π.
- $-5\sin t$ is the real part of $5e^{i(t+\frac{\pi}{2})} = 5e^{it}e^{i\pi/2}$ and has amplitude 5 and phase $\pi/2$.

Example: What is the amplitude and phase of:

$$\sin t - \cos t? \tag{41}$$

1. Convert each part into a complex form:

- $\sin t$ is the real part of $e^{i(t-\frac{\pi}{2})}$.
- $-\cos t$ is the real part of $e^{i(t+\pi)} = e^{it}e^{i\pi}$.

Thus

$$\sin t - \cos t = \Re\left(e^{i(t+-\frac{\pi}{2})} + e^{it}e^{i\pi}\right) = \Re\left(e^{it}\left[e^{-i\frac{\pi}{2}} + e^{i\pi}\right]\right)$$

so can be represented by the complex wave

$$e^{it}\left[e^{-i\frac{\pi}{2}} + e^{i\pi}\right].$$

2. Now compare to the complex wave solution of the form $Ae^{i(t+\phi)}$ (where $\omega = 1$ here). Dividing through by e^{it} gives

$$\begin{aligned} Ae^{i\phi} &= e^{-i\frac{\pi}{2}} + e^{i\pi} \\ \text{or} \quad -1-i &= Ae^{i\phi} \end{aligned}$$

Taking the modulus: $A = |-1-i| = \sqrt{2}$.

Real part : $\sqrt{2}\cos\phi = -1$, Imaginary part : $\sqrt{2}\sin\phi = -1$

So, $\cos\phi = -\dfrac{1}{\sqrt{2}}$, $\sin\phi = -\dfrac{1}{\sqrt{2}}$.

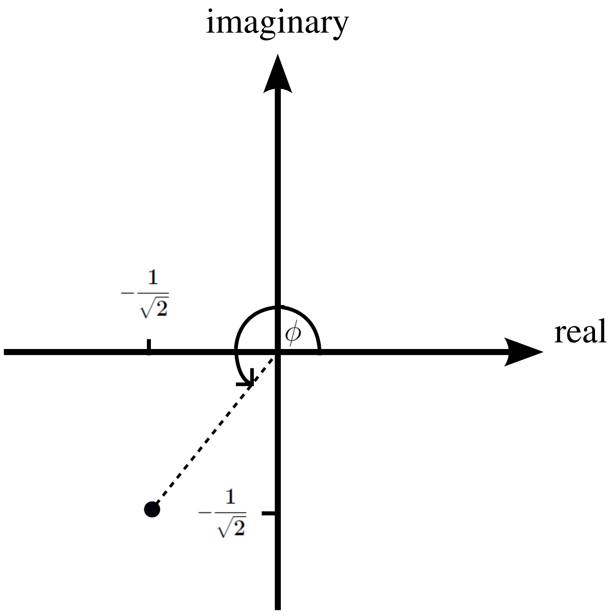

Figure 11

In Figure 11, $\phi = \pi + \frac{\pi}{4} = \frac{5\pi}{4}$.

NB — you need to take care with finding the phase ϕ, since the inverse sine and cosine functions may give the correct answer in the wrong range. i.e. your calculator will tell you that

$$\sin^{-1}\frac{-1}{\sqrt{2}} = \frac{-\pi}{4}$$

and

$$\tan^{-1} 1 = \frac{\pi}{4}.$$

These are correct, but you must interpret them correctly! To be sure, draw a picture.

4.10.2 Complex solution to the wave equation

A problem in time-series analysis might be to find the solution to the equation:

$$\frac{d^2y}{dt^2} = -y \;; \qquad y(0) = 2, y'(0) = 1.$$

The answer turns out to be:

$$y = \sin t + 2\cos t. \tag{42}$$

Equation 42 is an expression for a wave. But how might we discover this solution? One method is to attempt a trail solution of the form:

$$y = A\sin t + B\cos t,$$

and then try to find A and B.

Another is to find a "complex wave" solution of the form,

$$y = Ce^{it},$$

where C is a complex number and as stated at the beginning of this section, we are only interested in the real component, $\Re(y)$, as the physical solution. In this case, $C = 2 - i$, so that:

$$\begin{aligned}
\Re(y) &= \Re\left((2-i)\left(\cos t + i\sin t\right)\right), \\
&= \Re\left(2\cos t + \sin t + i\left(2\sin t - \cos t\right)\right), \\
&= 2\cos t + \sin t.
\end{aligned}$$

The advantage of this method is that we need only find a single unknown number C, rather than two (A and B).

5 Error analysis

⇒ **Wolfram link** (*http://mathworld.wolfram.com/Mean.html*)

⇒ **Wolfram link** (*http://mathworld.wolfram.com/StandardDeviation.html*)

⇒ **Wolfram link** (*http://mathworld.wolfram.com/RelativeError.html*)

⇒ **Wolfram link** (*http://mathworld.wolfram.com/AbsoluteError.html*)

Suppose we make an observation of a quantity x, called x_i, which we repeat n times to get the set: $\{x_1, x_2, x_3 ... x_n\}$.

The **mean**, μ, is:	$\frac{1}{n}\sum_{i=1}^{n} x_i$.
The **variance**, σ^2, is:	$\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 = \frac{1}{n}\left[\sum_{i=1}^{n} x_i^2\right] - \mu^2$.
The **standard deviation**, σ, is:	$\sqrt{\text{variance}} = \sqrt{\sigma^2}$.
The **mode** is:	the *most common value*.
The **median** is:	the *middle value* if the values of x_i are written in numerical order.

Suppose x_1 is an observation of the true value x.

The **absolute error** is: $x_1 - x$.

The **relative error** is: $\frac{x_1 - x}{x}$.

Accuracy is how close a measured value is to the true value (i.e. absolute error).

Precision is how close the (repeated) measure values are to each other.

We can have:

- high precision yet low accuracy,
- high accuracy yet low precision,
- high accuracy and high precision.

Accuracy and precision may differ if there is bias in the measurement. For example if a set of high precision digital scales read "1 kg" instead of "0" with nothing on the scales, then all measurements of weight are 1 kg out. Therefore, the scales may be precise yet not accurate.

5.1 Plus/Minus Notation

Errors are often quoted in the form: value $\pm$ standard deviation units.
For example: $\rho = 100 \pm 1$ kg/cm^3, where 1 here is the standard deviation. With this terminology, it is possible that ρ is equal to 98 or 103 kg/cm^3, although very unlikely.

Sometimes, the absolute range of values are given. The range of values of ρ may be:
$99 \leq \rho \leq 101 \mathrm{kg\,cm^{-3}}$. With this terminology, it is NOT possible for ρ to be $98\,\mathrm{kg\,cm^{-3}}$.

5.2 Propagation of errors

$\Rightarrow$ **Wolfram link** (*http://mathworld.wolfram.com/ErrorPropagation.html*)

Suppose we estimate the mass of a sphere using the formula: $m = \frac{4\pi}{3}\rho r^3$, where r and ρ have a measurement error. What is the error in m?

We can use the formula:

$$\delta m = \sqrt{\left(\frac{\partial m}{\partial r}\right)^2 (\delta r)^2 + \left(\frac{\partial m}{\partial \rho}\right)^2 (\delta \rho)^2}\ , \tag{43}$$

where δm is the error in mass, m.

e.g. if $r = 0.1 \pm 0.01$ m and $\rho = 100 \pm 1\ \mathrm{kg\,m^{-3}}$, then

$$m = \frac{4\pi}{3} \times 100 \times 0.1^3\ ,$$
$$\delta m = \sqrt{(4\pi\rho r^2 \times 0.01)^2 + \left(\frac{4\pi}{3} r^3 \times 1\right)^2}\ .$$

Hence, $m = 0.419 \pm 0.004$ kg i.e. $\sim 1\%$ relative error.

This formula extends to any number of variables. If $f = f(x, y, z)$, then,

$$\delta f = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 (\delta x)^2 + \left(\frac{\partial f}{\partial y}\right)^2 (\delta y)^2 + \left(\frac{\partial f}{\partial z}\right)^2 (\delta z)^2}\ . \tag{44}$$

Note: δf or δm gives the standard deviation of f or m, calculated as a function of the standard deviation of its dependant variables.

What about adding quantities? For example, if $f = x + y$, where both x and y have an error, what is the error in f?

This is actually a special case of the general formula (equation 44):

$$\delta f = \sqrt{(\delta x)^2 + (\delta y)^2}, \qquad \text{since } \frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 1 \ . \tag{45}$$

5.3 Comparison with "worst case" scenario?

An alternative to using equation 44 is to compute the worst case error. For example, if the density measurement was incorrect by a single standard deviation above its mean, and the radius too high by a single standard deviation, then the mass calculated would be 0.563kg, compared to the mean value of 0.419kg. This is 0.14kg above the mean, or $\sim 30\%$, a lot higher than that calculated above. But this is not a fair calculation, as it is very unlikely that both the density and radius conspire together in this way.

5.4 Normal Distribution

If a variable x is normally distributed with mean μ and standard deviation, σ, it has a probability-density function (or likelihood function) as shown in Figure 12.

The most probable value of x is μ. The probability of other values are also known:
$\mu - \sigma \leq x \leq \mu + \sigma$: occurs with a probability of 68%.
$\mu - 2\sigma \leq x \leq \mu + 2\sigma$: occurs with a probability of 95%.
$\mu - 3\sigma \leq x \leq \mu + 3\sigma$: occurs with a probability of 99.7%.

In Geophysics, many variables are assumed to be normally distributed. The formula that

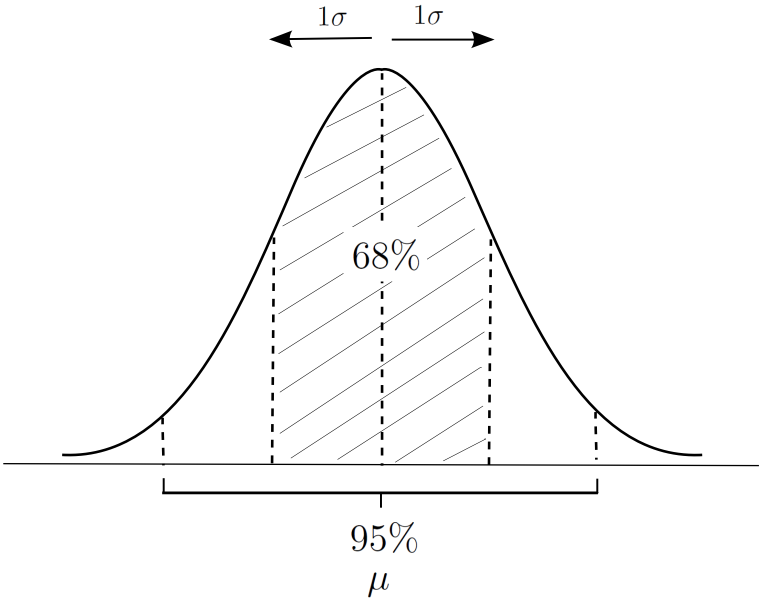

Figure 12

gives the likelihood of x is:

$$f(x) \quad = \quad \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{46}$$

which is normalised such that

$$\int_{-\infty}^{\infty} f(x) \; dx \quad = \quad 1. \tag{47}$$

5.5 Central limit theorem

Normal distributions often arise out of other non-normal distributions. If x_1, x_2, etc ... are identical variables then we can define $X = x_1 + x_2 + \cdots + x_M$.

It turns out that X is approximately normally distributed with mean $M\mu$ and standard deviation $\sqrt{M}\sigma$, where each of the x_i has mean μ and standard deviation σ, i.e.

$$X \sim N\left(M\mu, \sqrt{M}\sigma\right).$$

For example: a die shows a random number from 1 to 6 with equal probability for each.

The mean of the value shown by a single die x_i is $\mu = \frac{1}{6}(1+2+3+4+5+6) = 3.5$ and the standard deviation is found by $\sigma = \sqrt{\frac{1}{6}\left(1^2+2^2+3^2+4^2+5^2+6^2\right) - \mu^2} = 1.71$.

A single die is <u>not</u> normally distributed, however, if we throw 10 dice and add their scores, the totals $X = x_1 + x_2 + \cdots + x_{10}$ will be approximately normally distributed. For 10 dice the mean is $10\times\mu = 10\times3.5 = 35$, and the standard deviation is $\sqrt{10}\times\sigma = \sqrt{10}\times1.71 = 5.37$.

When considering the average of the ten dice, i.e. $Y = (x_1 + ... + x_{10})/10$, then Y has mean μ, the same mean as each of the x_i. The standard deviation of $10Y$ is $\sqrt{M}\sigma$, so the standard deviation of Y is $\frac{1}{10}\sqrt{10}\sigma = \frac{\sigma}{\sqrt{10}}$. This occurs because the standard deviation is linear in any factor applied to the variables (here, the factor $1/10$).

5.6 Confidence Intervals

Suppose a variable X is distributed normally with mean μ and standard deviation σ. Then the 68% confidence interval for X is $[\mu-\sigma, \mu+\sigma]$, as there is a 68% probability of X taking a value in this range. Similarly the 95% confidence interval is (approximately) $[\mu-2\sigma, \mu+2\sigma]$.

<u>For example:</u> Suppose $x_1, x_2, \ldots, x_{25}$ are measurements all with mean 10 g and standard deviation 2 g. What is the 95% confidence interval for $X = x_1 + x_2 + \cdots + x_{25}$?

Solution:
X is distributed normally with mean 250 g and standard deviation $\sqrt{25}\times 2 = 10$ g. The 95% confidence interval is then $[230, 270]$ g of the form $[\mu-2\sigma, \mu+2\sigma]$.

Additional material:
⇒ **<u>Video link</u>** (*http://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2005/*)