Vectors & Matrices

Introduction

Vectors were developed from Quaternions (which are really cool) in the latter part of the 19th century, quaternions themselves being generalisation of complex numbers (square roots of negative numbers) developed in the 16th century. Vectors are particularly useful in engineering and architecture for dealing with geometry in three-dimensional space.
Matrices were the idea of Arthur Cayley, a lawyer who became a mathematician. In 1858 he demonstrated the use of matrices in the solution of simultaneous equations, a purpose for which they are widely used today. Many types of problems in engineering give rise to linear algebraic simultaneous equations, and matrices are also useful for analysing complex linear structures by computational or numerical methods.

Outline:

Vectors & Matrices

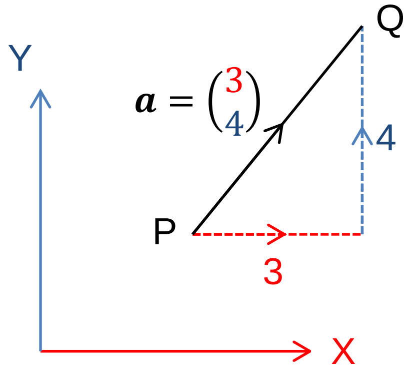
Whilst everything which is discussed in this introduction works for vectors and matrices in three, four and higher dimensions, these notes generally explain everything in 2D first, in order to make the notes less complicated to read and easier to understand. Later we are going to use it all for some very complicated work in 3D, and the tutorial questions will involve you using these concepts in 3D, but for now we will keep things simple.

Vector Algebra

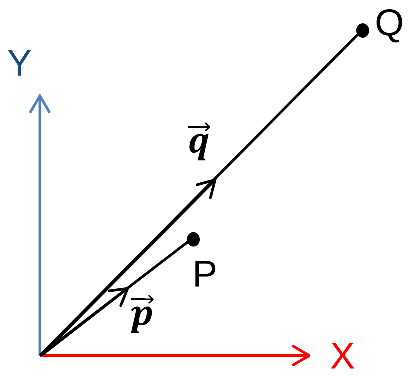
In its simplest form, a vector is just a list of quantities. However, in engineering, it is usual to use a geometrical interpretation of vectors, where the separate quantities relate to amounts of translation along a set of axes. Vectors therefore have a *direction* and *magnitude*. They are usually written as a column of numbers or "elements", and are represented in algebra with a lower-case letter in bold, or underlined, or with a little arrow above, such as $\boldsymbol{a}$, $\underline{a}$ or $\vec{a}$.

For example, in 2D, if we let the first quantity represent the x-displacement and the second the y-displacement, then the vector can represent any given displacement in the XY plane:

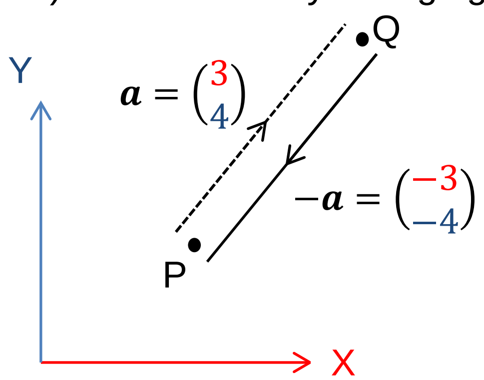
The vector **a** below represents a displacement from point P to point Q, which is 3 units along the X-axis and 4 units along the y-axis. Its direction is from point P towards point Q.



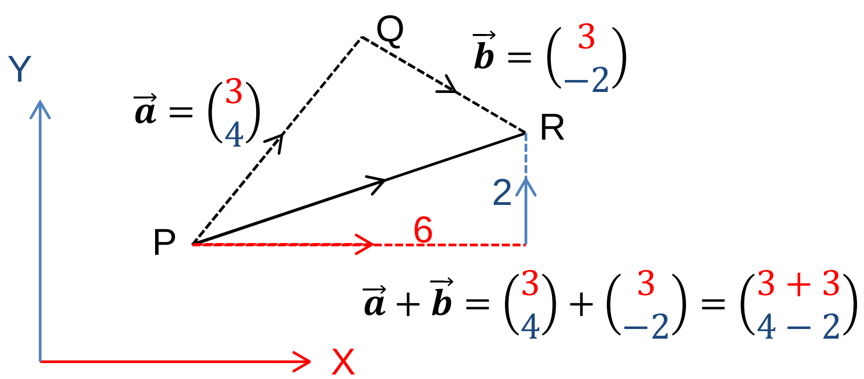
Vectors can also be used to represent points in space, by defining the displacement of the vector to be relative to the coordinate system origin (0,0). The sketch below for example shows position vectors **p** and **q** representing points P and Q respectively.



If a vector **a** represents a displacement from point P to point Q, then the negative vector **–a** has the same length but the opposite direction (i.e. from point Q to point P) and is found by changing the sign of each element in turn.

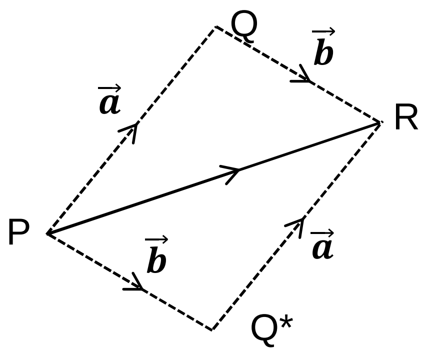


Vectors can be added together, and this can be thought of as first making one displacement and then another. For example, if a vector **a** represents a displacement from point P to point Q, and similarly a vector **b** represents a displacement from point Q to point R, then the sum **a+b** can be thought of as a displacement from P to Q and then from Q to R, which is the same as going from P to R directly.



Addition

The sum of two vectors is found algebraically by summing the individual parts, as shown above. This obviously means that only vectors of the same dimension (number of elements) can be added. It also follows that the order of vector addition is not important, since one could move from P to R by either first moving in direction **a** and then **b** (via point Q), or first **b** and then **a** (via point Q* as shown below).



$$\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}} = \vec{\boldsymbol{b}} + \vec{\boldsymbol{a}}$$

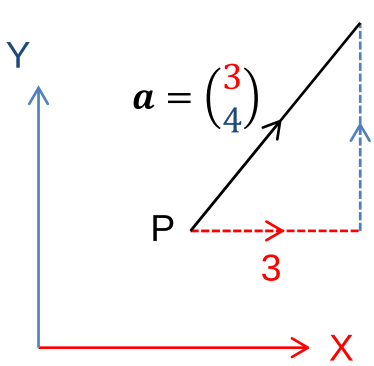
Multiplication by a scalar

From this, and the usual mathematical definition of multiplication being the same as adding something to itself a number of times, vectors can be scaled by a constant in the obvious way. For example:

$$\vec{\boldsymbol{a}} + \vec{\boldsymbol{a}} + \vec{\boldsymbol{a}} = \begin{pmatrix} 3 \\ 4 \end{pmatrix} + \begin{pmatrix} 3 \\ 4 \end{pmatrix} + \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 3+3+3 \\ 4+4+4 \end{pmatrix} = \begin{pmatrix} 9 \\ 12 \end{pmatrix} = \begin{pmatrix} 3 \times 3 \\ 3 \times 4 \end{pmatrix} = 3 \times \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \boldsymbol{3\vec{a}}$$

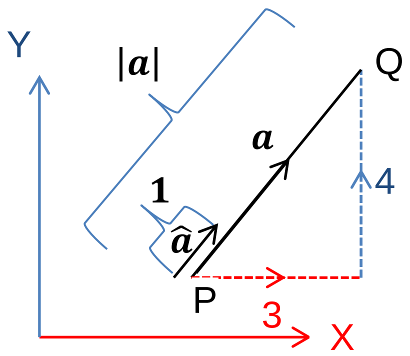
Magnitude

The distance between points P and Q can be found using Pythagoras as $= \sqrt{3^2 + 4^2} = 5$, where "3" and "4" are the x- and y- displacements respectively, and are therefore the elements of the displacement vector.



In this way, the length of a vector (also known as its *magnitude* or *modulus*) can be found using Pythagoras by taking the square-root of the sum of the squares of the elements. The modulus of a vector is denoted in writing by placing it between vertical lines, so the modulus of **a** is written as $|\boldsymbol{a}|$ or sometimes just $|a|$ without the bold letter, since it is implicit that the lower-case “a” inside the lines represents a vector.

Unit Vector
Combining the two previous assumptions, we can scale any vector such that it has unit length, by finding its length “L” and then multiplying the vector by “1/L”. A vector with length=1 is said to be a “unit vector” and is usually denoted with a little “^” (hat) on top.

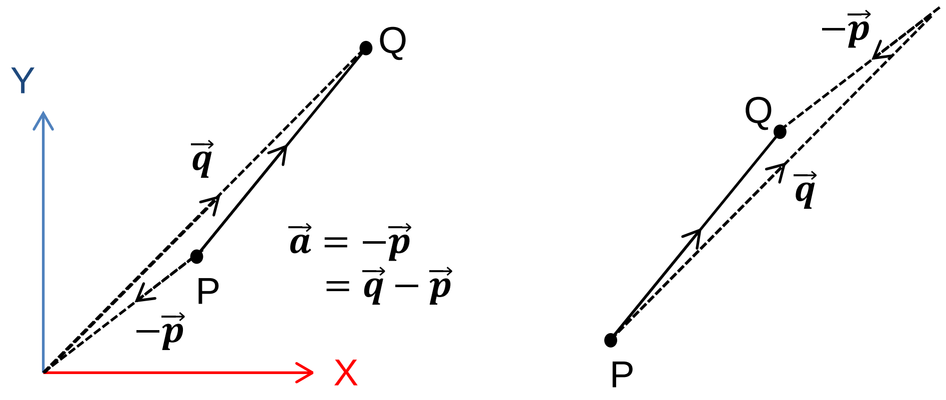


$$\boldsymbol{a} = \begin{pmatrix} 3 \\ 4 \end{pmatrix} \quad \Rightarrow \quad |\boldsymbol{a}| = \sqrt{3^2 + 4^2} = 5 \quad \Rightarrow \quad \hat{\boldsymbol{a}} = \frac{\boldsymbol{a}}{|\boldsymbol{a}|} = \frac{1}{5}\begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix}$$

Subtraction
Vectors can also be subtracted by defining this as the addition of the negative, just like in algebra:

a – **b** = **a** + (-**b**)

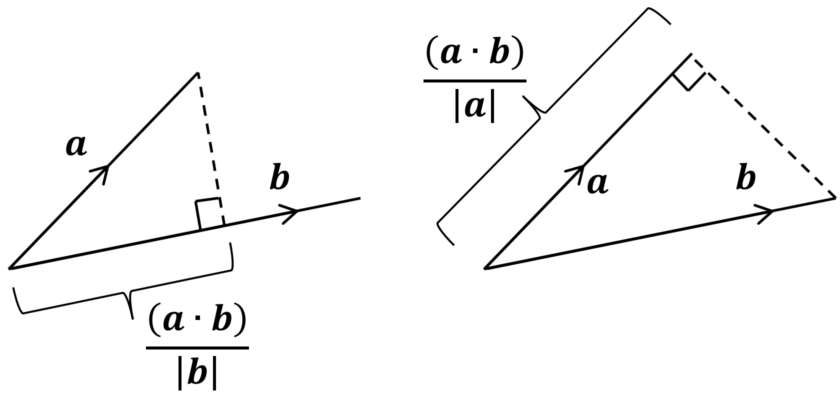
From the above, it follows that we can express the vector **a**, which represents a displacement from point P to point Q as the difference of the coordinates.



Scalar (Dot) Product

The scalar product is an operator (like + and -) which takes two vectors and turns them into a single scalar quantity. It is written as a dot between the two vectors (hence the alternative name *dot product*). So the scalar product of **a** and **b**, say, is $\boldsymbol{a} \cdot \boldsymbol{b}$ and represents the size of one vector × the size of the projection of the other vector onto it.

a·b = |a| × |projection of b onto a|
= |b| × |projection of a onto b|



It can be calculated by summing the products of the corresponding elements as follows:

$$a \cdot b = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1 \times b_1 + a_2 \times b_2 + \cdots + a_n \times b_n$$

And it can be shown that:

$$a \cdot b = |a||b| \cos\theta$$

where θ is the angle between the two vectors.

It follows directly that:

$a \cdot b = b \cdot a$ (i.e. it is *commutative*)

$a \cdot b = 0$ if the two vectors are perpendicular

$a \cdot a = |a|^2$ since vector **a** is parallel to itself, so θ=0

The dot product is *distributive* over addition and subtraction, which means we can multiply out brackets as:

$$(a \pm b) \cdot c = a \cdot c \pm b \cdot c$$

Example: If $a = \begin{pmatrix} 2 \\ 3 \\ -6 \end{pmatrix}$ and $b = \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}$ find the angle between the two.

$$|a| = \sqrt{2^2 + 3^2 + (-6)^2} = 7$$
$$|b| = \sqrt{3^2 + 0^2 + 4^2} = 5$$

$$a \cdot b = 2 \times 3 + 3 \times 0 + (-6) \times 4 = 6 + 0 - 24 = -18$$

$$\theta = \cos^{-1}\left(\frac{a \cdot b}{|a||b|}\right) = \cos^{-1}\left(\frac{-18}{7 \times 5}\right) = 2.11rad$$

Matrix Algebra

A matrix is a grid of numbers arranged in a bracketed rectangular array of rows and columns. They are usually represented with upper case letters, either in bold text, underlined or in square brackets, such as **A**, $\underline{A}$ or [A]. For example, a matrix **A** might be:

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

A has *m* rows and *n* columns. The order of **A** is ($m \times n$). If $n=m$ then the matrix is *square* of order *n*. The individual quantities a_{ij} making up the matrix are known as *elements* of the matrix. Two matrices are equal if and only if they have the same order and all the corresponding elements are equal.

A matrix with a single column can be thought of as a vector, and much of the algebra derived above also follows for matrices. For completeness we will briefly look at elementary algebra here.

Addition

To add two matrices we simply add together each corresponding *element*.

If **C** = **A** + **B**, then $c_{ij} = a_{ij} + b_{ij}$, where i=row, j=column

We can only do this if the two matrices *have the same order*.

Example:

$$\begin{bmatrix} 5 & 5 & -2 \\ 1 & 4 & 1 \end{bmatrix} + \begin{bmatrix} 4 & 0 & -9 \\ -1 & 8 & 5 \end{bmatrix} = \begin{bmatrix} 9 & 5 & -11 \\ 0 & 12 & 6 \end{bmatrix}$$

As with vectors, it is clear that matrix addition is *commutative*:

A + **B** = **B** + **A**

and *associative*:

A + (**B** + **C**) = (**A** + **B**) + **C**

Subtraction

Subtraction is defined in the same way, i.e.:

If **C** = **A** - **B**, then $c_{ij} = a_{ij} - b_{ij}$, where i=row, j=column

Example:

$$\begin{bmatrix} 5 & 5 & -2 \\ 1 & 4 & 1 \end{bmatrix} - \begin{bmatrix} 4 & 0 & -9 \\ -1 & 8 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 5 & 7 \\ 2 & -4 & -5 \end{bmatrix}$$

Multiplication by a scalar

Just as with vectors, multiplying a matrix by a scalar simply multiplies *each element* by the scalar:

If $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$, then $3A = \begin{bmatrix} 3a_{11} & 3a_{12} & 3a_{13} \\ 3a_{21} & 3a_{22} & 3a_{23} \end{bmatrix}$

Matrix Multiplication

We can form the product **AB** only if the number of *columns* of **A** equals the number of *rows* of **B**. If this is the case the matrices are said to be *conformable*.

To form the first element (position 1,1) of **AB**, we must take the dot-product of the elements of the first *row* of **A** and the first *column* of **B**. Similarly the element of **AB** in the (1,2) position is given by "dot-producting" the first *row* of **A** by the second *column* of **B**, and so on. e.g.:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$$
$$= \begin{bmatrix} (a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31}) & (a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32}) \\ (a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31}) & (a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32}) \end{bmatrix}$$

$$\begin{bmatrix} \cdots & A_{Row1} & \cdots \\ \cdots & A_{Row2} & \cdots \end{bmatrix} \begin{bmatrix} \vdots & \vdots \\ B_{Col1} & B_{Col2} \\ \vdots & \vdots \end{bmatrix}$$
$$= \begin{bmatrix} (A_{Row1} \cdot B_{Col1}) & (A_{Row1} \cdot B_{Col2}) \\ (A_{Row2} \cdot B_{Col1}) & (A_{Row2} \cdot B_{Col2}) \end{bmatrix}$$

More formally, if **C = AB** where **A** has order (m×n) and **B** has order (n×p) then an element of **C** is given by:

$$C_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

And the result, **C** has the order (m×p) i.e. m rows and p columns.

Note that, in general **AB≠BA**

Example: If **A**= $\begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix}$ and **B**= $\begin{bmatrix} 3 & 1 \\ 2 & 3 \\ 1 & 2 \end{bmatrix}$ find **AB**

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

Transpose

The transpose of a matrix **A** is denoted by $\mathbf{A}^{T}$ and is formed by *swapping rows* and *columns.* The row become the columns and the columns become the rows, in effect mirroring the matrix along the leading diagonal. So, if:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \quad \text{then} \quad A^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$$

Special Types of Matrices

Row and column vectors

A column vector is a matrix of order (m×1). Column vectors are usually denoted as vectors by lower case bold letters (in typescript) or underlined lowercase letter (in handwriting):

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \quad \text{and} \quad \boldsymbol{a}^T = [a_1 \quad a_2 \quad \cdots \quad a_m]$$

The transpose of a column vector is a *row vector* of order (1×m).

Zero (or null) matrix

An (m×n) matrix where *all* the elements are *zero* is known as a *zero* or *null matrix* of order (m×n), often written $\mathbf{O}_{mn}$, e.g.:

$$\boldsymbol{O_{32}} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Clearly, as with the number zero in algebra, adding the appropriate order zero matrix to any matrix **A** leaves it unchanged:

$$\mathbf{A} + \mathbf{O} = \mathbf{O} + \mathbf{A} = \mathbf{A}$$

Diagonal Matrices

Diagonal matrices are *always square* and have all elements equal to zero *except those on the leading diagonal*, i.e.:

$a_{ij} = 0$ if $i \neq j$ e.g. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

Identity matrix

The *identity matrix* (or unit matrix) is very important in matrix algebra. It is a *square* matrix of order ($m\times m$), denoted by **I**, and is a diagonal matrix with all the elements on the leading diagonal equal to 1, i.e.

$a_{ij} = 0$ if $i \neq j$

$a_{ij} = 1$ if $i = j$

e.g. $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Under multiplication, as with the number one in algebra, multiplying any ($m\times m$) matrix **A** by **I** leaves **A** unchanged:

$$\mathbf{A}\,\mathbf{I} = \mathbf{I}\,\mathbf{A} = \mathbf{A}$$

More generally, if **A** is an ($m\times n$) matrix, then:

$$\mathbf{I}_m \mathbf{A} = \mathbf{A} \quad \text{and} \quad \mathbf{A}\,\mathbf{I}_n = \mathbf{A}$$

Representation of Linear Equations

Matrix multiplication is useful for representing systems of linear simultaneous equations. We will come back to this later, but for now note for example:

$$\begin{aligned} x + y + z &= 5 \\ 2x + 4y + z &= 3 \\ 6y + z &= 1 \end{aligned}$$

can be written:

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 4 & 1 \\ 0 & 6 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix}$$

If you don't believe me, multiply out the left-hand-side…

Determinants

Analogous to the modulus of a vector, a determinant somehow gives a single measure of the collective size of the elements in a square matrix. Again, similar to the modulus of a vector, the determinant is denoted by placing the matrix within vertical lines, thus the determinant of a matrix **A** is written $|\boldsymbol{A}|$, and again, the bold typeface for the matrix is sometimes ignored, as $|A|$, since it is obvious from the capital letter and vertical lines that the thing inside is a matrix. Occasionally the determinant is written as a function, det(A).

First Order
In the most trivial of cases, a first order matrix is a 1x1 matrix containing a single number. In this case, the obvious way of measuring the size of the element is to use that single element. So:

$$\text{If } A = [\mathrm{a}_{11}] \text{ then } |A| = |\mathrm{a}_{11}| = \mathrm{a}_{11}$$

For example:

$$\text{If } A = [4] \text{ then } |A| = |4| = 4$$

Second Order
More generally, determinants arise naturally from the solution of a set of linear simultaneous equations. Suppose we have the following pair of equations where x_1 and x_2 are our unknowns:

$$a_{11}x_1 + a_{12}x_2 = b_1$$
$$a_{21}x_1 + a_{22}x_2 = b_2$$

Solving for x_1 and x_2 in the usual way, we find that:

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}} \quad \text{and} \quad x_2 = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}}$$

Provided $a_{11}a_{22} - a_{12}a_{21} \neq 0$. We now define this quantity as:

$$\mathrm{D} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \equiv a_{11}a_{22} - a_{12}a_{21}$$

We call this a *determinant*. In this case, since there are two columns and two rows, it is a second order determinant. The diagonal containing a_{11} and a_{22} is called the *leading diagonal*.

The second-order determinant is given by multiplying the two numbers in the leading diagonal and subtracting the product of the numbers in the other diagonal.

Example:

$$\begin{vmatrix}1 & -4\\ 2 & 3\end{vmatrix} = \begin{vmatrix}1 & \square\\ \square & 3\end{vmatrix} - \begin{vmatrix}\square & -4\\ 2 & \square\end{vmatrix} = 1 \times 3 - (-4) \times 2 = 11$$

Third Order

If we now consider the system of equations:

$$a_{11}x + a_{12}y + a_{13}z = b_1$$
$$a_{21}x + a_{22}y + a_{23}z = b_3$$
$$a_{31}x + a_{32}y + a_{33}z = b_3$$

The *denominator* of the solutions (i.e. expressions for *x*, *y* and *z*) is defined as a 3×3 determinant:

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

Again, this can be found by looking at the multiples of the diagonal entries. However, we need to “wrap around”:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & (a_{11}) & (a_{12})\\ a_{21} & a_{22} & a_{23} & (a_{21}) & (a_{22})\\ a_{31} & a_{32} & a_{33} & (a_{31}) & (a_{32})\end{vmatrix}$$

i.e.:

$$D = +a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

$$= +a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

It can now be seen that the determinant can be expressed in terms of 2×2 determinants:

$$\mathrm{D}_3 = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$
$$= a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13}$$

Here, M_{11}, M_{12} and M_{13} are the second order determinants remaining when the row and column containing a_{11}, a_{12} and a_{13} respectively, are deleted, e.g.:

$$\mathrm{M}_{11} = \begin{vmatrix} a_{11} & \times & \times \\ \times & \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} \\ \times & \end{vmatrix}$$

These determinants are called the *minors* of a_{11}, a_{12} and a_{13}. Similarly we could have expressed the determinant in terms of a_{21}, a_{22} and a_{23} with corresponding minors, M_{21}, M_{22} and M_{23} where, for example, the minor M_{21}, corresponding to a_{21} would be formed as:

$$\mathrm{M}_{21} = \begin{vmatrix} \times & \begin{vmatrix} a_{12} & a_{13} \\ \times & \times \\ a_{32} & a_{33} \end{vmatrix} \\ a_{21} & \\ \times & \end{vmatrix}$$

So, in this case the determinant is:

$$\mathrm{D}_3 = -a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{22}\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{23}\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$$
$$= -a_{21}M_{21} + a_{22}M_{22} - a_{23}M_{23}$$

The result would be exactly the same. In fact we can expand in terms of *any* row *or* column. So, in general the minor, M_{ij}, is given by removing row *i* and column *j* from the matrix and forming the determinant of what remains.

Note, however, that the signs of each *product* change according to the rule of alternating signs:

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

If we define the *cofactor*, c_{ij} of a_{ij} as: $C_{ij} = (-1)^{i+j}\, M_{ij}$ then the determinant becomes:

$$D = a_{11}c_{11} + a_{12}c_{12} + a_{13}c_{13}$$

or, in general: $D = \sum_{j=1}^{n} a_{ij}c_{ij}$ when expanding along row i.

Example: Evaluate the determinant: $\begin{vmatrix} 4 & -2 & 3 \\ -3 & 0 & -1 \\ 1 & -4 & 2 \end{vmatrix}$ using a) the 1st row, and b) the 3rd column

a)

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

$= 4\times(0\times2-(-1)\times(-4)) + 2\times((-3)\times2-(-1)\times1) + 3\times((-3)\times(-4)-0\times1)$
$= 4\times(0-4) \quad + 2 \times (-6+1) \quad + 3 \times (12-0)$
$= -16 \quad - 10 \quad + 36$
$= 10$

b)

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

$= 3\times((-3)\times(-4)-0\times1) + 1\times(4\times(-4)-(-2)\times1) + 2\times(4\times0-(-2)\times(-3))$
$= 3\times(12-0) \quad + 1\times(-16+2) \quad + 2\times(0-6)$
$= 36 \quad - 14 \quad - 12$
$= 10$

An unexpected property of determinants is that when any row is multiplied by the cofactors of a *different* row, the result is *always zero*. For example, consider the expansion:

$$a_{11}c_{21}+a_{12}c_{22}+a_{13}c_{23} = 0$$

Here, the *a*'s come from row 1 but the cofactors, *c*, are associated with row 2. So, expanding the expression where:

$$\mathrm{D} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

We get:

$= a_{11}c_{21}+a_{12}c_{22}+a_{13}c_{23}$
$= -a_{11}(a_{12}a_{33}-a_{13}a_{32}) + a_{12}(a_{11}a_{33}-a_{13}a_{31}) - a_{13}(a_{11}a_{32}-a_{12}a_{31})$
$= -a_{11}a_{12}a_{33}+a_{11}a_{13}a_{32}+a_{12}a_{11}a_{33}-a_{12}a_{13}a_{31}-a_{13}a_{11}a_{32}+a_{13}a_{12}a_{31}$
$= 0$

So, in general:

$$\sum_{k=1}^{n} a_{ik}c_{jk} = 0 \quad i \neq j$$
$$= D \quad i = j$$

Properties of Determinants

There are several properties of determinants which we can use to help simplify the calculation of the determinant. In the following, we will refer to:

$$\mathrm{D} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

If rows and columns are interchanged the value of the determinant remains the same:

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

i.e. $|\mathrm{A}| = |\mathrm{A}^{\mathrm{T}}|$

If two rows (or columns) are exchanged, the determinant changes sign:

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

If two rows (or columns) are identical then the determinant is zero. This follows from the previous property, i.e. exchanging rows changes the sign of the determinant, then if two identical rows are exchanged this can only be the case if the determinate is zero.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{vmatrix} = 0$$

If all the elements of any row (or column) are multiplied by a common factor, the value of the determinant is multiplied by this factor:

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

If any two rows (or columns) are proportional to each other, then it follows from the previous two properties that the determinant is zero.

$$\begin{vmatrix} a_{11} & \lambda a_{11} & a_{13} \\ a_{21} & \lambda a_{21} & a_{23} \\ a_{31} & \lambda a_{31} & a_{33} \end{vmatrix} = \lambda \begin{vmatrix} a_{11} & a_{11} & a_{13} \\ a_{21} & a_{21} & a_{23} \\ a_{31} & a_{31} & a_{33} \end{vmatrix} = \lambda 0 = 0$$

If a multiple of a row (or column) is added to another row (or column) the value of the determinant remains *unchanged*:

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

If the matrix is *diagonal* then $D = a_{11}a_{22}a_{33}a_{44}\ldots a_{nn}$ i.e. expand along first row:

$$D = \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & 0 \\ 0 & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33}$$

If the determinant is *triangular* then $D=a_{11}a_{22}a_{33}a_{44}\ldots a_{nn}$ i.e. expand down first column:

$$D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ 0 & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33}$$

We will see the significance of some of these properties later!

Higher Order Determinants

An n^{th} order determinant can be evaluated in exactly the same way as for the 3rd order determinant:

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1\mathrm{n}} \\ a_{21} & a_{22} & \cdots & a_{2\mathrm{n}} \\ \vdots & \vdots & \ddots & \vdots \\ a_{\mathrm{n}1} & a_{\mathrm{n}2} & \cdots & a_{\mathrm{nn}} \end{vmatrix} = \sum_{j=1}^{n} a_{\mathrm{ij}} c_{\mathrm{ij}}$$

The cofactors are defined in the usual way: $c_{ij} = (-1)^{i+j} M_{ij}$

In fact the method of expansion along a row / column by taking alternating signs, deleting the row & column and multiplying by determinant of the rest is completely *recursive*.

$$|a_{11}| = a_{11}$$

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} +a_{11} & \times \\ \times & |a_{22}| \end{vmatrix} + \begin{vmatrix} \times & -a_{12} \\ |a_{21}| & \times \end{vmatrix}$$

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

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

However, calculating the minors of higher order determinants can become quite tedious by hand.

For example, evaluate the following 4th order determinant:

$$\begin{vmatrix} 2 & 3 & 5 & -1 \\ 1 & 2 & -3 & 2 \\ 4 & -1 & 2 & 5 \\ 2 & -2 & 3 & 1 \end{vmatrix}$$

We could expand by the 1st row:

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

$$= 2\left[\; 2\begin{vmatrix} 2 & 5 \\ 3 & 1 \end{vmatrix} - (-3)\begin{vmatrix} -1 & 5 \\ -2 & 1 \end{vmatrix} + 2\begin{vmatrix} -1 & 2 \\ -2 & 3 \end{vmatrix} \;\right] - 3\left[\qquad\right] + 5\left[\qquad\right] \ldots\text{etc}$$

$$= 2\left[\; 2(2|1| - 5|3|) - (-3)(-1|1| - 5|-2|) + 2(-1|3| - 2|-2|) \;\right] - 3[\;\;] + 5[\;\;] \ldots\text{etc}$$

$$= 2[\; -26 + 27 + 2 \;] - 3[-13 - 18 + 16] + 5[9 + 12 - 12]$$

= 6 + 45 + 45 + 3 = 99

However, using the properties discussed earlier, we can be much more cunning…

For example, subtracting multiples of one row from another does not change the determinant. So, since row two has a "1" in the first column, we can subtract multiples of row two from the others to make their first entries zero:

$$\begin{vmatrix} 2 & 3 & 5 & -1 \\ 1 & 2 & -3 & 2 \\ 4 & -1 & 2 & 5 \\ 2 & -2 & 3 & 1 \end{vmatrix} = \begin{vmatrix} 0 & -1 & 11 & -5 \\ 1 & 2 & -3 & 2 \\ 0 & -9 & 14 & -3 \\ 0 & -6 & 9 & -3 \end{vmatrix} \begin{array}{l} \Leftarrow -2 \times \text{Row2} \\ \\ \Leftarrow -4 \times \text{Row2} \\ \Leftarrow -2 \times \text{Row2} \end{array}$$

So now we just need to expand down the first column and the problem is reduced to a 3x3 determinant (remembering the template of alternating signs:

$$= -1 \times \begin{vmatrix} -1 & 11 & -5 \\ -9 & 14 & -3 \\ -6 & 9 & -3 \end{vmatrix}$$

We can use the same trick with this 3x3 determinant, for example subtracting multiples of row one to give zeros in the first column:

$$= -\begin{vmatrix} -1 & 11 & -5 \\ -9 & 14 & -3 \\ -6 & 9 & -3 \end{vmatrix} = -\begin{vmatrix} -1 & 11 & -5 \\ 0 & -85 & 42 \\ 0 & -57 & 27 \end{vmatrix} \begin{array}{l} \\ \Leftarrow -9 \times \text{Row1} \\ \Leftarrow -6 \times \text{Row1} \end{array}$$

Expanding down first column therefore gives:

$$= -(-1) \times \begin{vmatrix} -85 & 42 \\ -57 & 27 \end{vmatrix}$$
$$= -85 \times 27 - 42 \times (-57)$$
$$= 99$$

Inverse of a Matrix

The inverse of a matrix is another extremely useful part of matrix algebra. The *inverse* $\mathbf{A}^{-1}$ of a matrix **A** is such that:

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

Only *square* matrices have an inverse. It is analogous to the reciprocal in algebra, where any number can be multiplied by its reciprocal to give one.

So, how do we find the inverse of a matrix? As an example we will examine a (*2*×*2*) matrix:

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

Suppose the *inverse* of this matrix is:

$$A^{-1} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$$

We want to find the values of *e, f, g* and *h* which satisfy the following:

$$AA^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Evaluating this (i.e. multiplying the two matrices) we get four simultaneous equations:

ae+bg = 1
af+bh = 0
ce+dg = 0
cf + dh = 1

Solving for the four unknowns we find:

$$e = \frac{d}{ad - bc} \qquad f = \frac{-b}{ad - bc}$$

$$g = \frac{-c}{ad - bc} \qquad h = \frac{d}{ad - bc}$$

Thus, we can say that:

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

Now, recalling the definition of *determinants*, the quantity *ad - bc* is the *determinant* of the (*2*×*2*) matrix **A**.

$$|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

So, to find the inverse of a (*2*×*2*) matrix **A:**

- Swap the elements on the leading diagonal.
- Change the sign of the other two elements.
- Divide all the elements by the determinant of **A**. This can be thought of as multiplying by 1/**|A|**.

The matrix: $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ is known as the *adjoint* of **A** or *adj*(**A**).

So, we can write the inverse as:

$$A^{-1} = \frac{\text{adj}(A)}{\det(A)}$$

If (*ad* – *bc*) = *0* then $\mathbf{A}^{-1}$ is undefined and the matrix **A** is said to be *singular*.

Example: Find the inverse of the following (*2*×*2*) matrix: $A = \begin{bmatrix} 3 & 7 \\ -5 & 2 \end{bmatrix}$

$$|A| = ad - bc = 3 \times 2 - 7 \times (-5) = 6 + 35 = 41$$

$$adj(A) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} 2 & -7 \\ 5 & 3 \end{bmatrix}$$

$$\Rightarrow A^{-1} = \frac{1}{41}\begin{bmatrix} 2 & -7 \\ 5 & 3 \end{bmatrix}$$

Adjoint Matrix

The inverse of larger square matrices can be found in exactly the same way as for the two by two matrix, i.e.

$$A^{-1} = \frac{\mathrm{adj}(A)}{\det(A)}$$

We know how to calculate the determinant of larger matrices, but how to we find the adjoint of **A**, *adj*(**A**)?

The matrix *adj*(**A**) is formed by assembling the matrix of the *cofactors* of **A** and taking its *transpose*. So for *square* matrix **A**:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\mathrm{n}} \\ a_{21} & a_{22} & \cdots & a_{2\mathrm{n}} \\ \vdots & \vdots & \ddots & \vdots \\ a_{\mathrm{n}1} & a_{\mathrm{n}2} & \cdots & a_{\mathrm{nn}} \end{bmatrix}$$

Recalling the definition of a cofactor: $c_{ij} = (-1)^{i+j} M_{ij}$

The matrix of cofactors, **C**, is:

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1\mathrm{n}} \\ c_{21} & c_{22} & \cdots & c_{2\mathrm{n}} \\ \vdots & \vdots & \ddots & \vdots \\ c_{\mathrm{n}1} & c_{\mathrm{n}2} & \cdots & c_{\mathrm{nn}} \end{bmatrix}$$

So that the adjoint of **A** is given by the transpose of the cofactor matrix:

$$adj(A) = C^T = \begin{bmatrix} c_{11} & c_{21} & \cdots & c_{\mathrm{n}1} \\ c_{12} & c_{22} & \cdots & c_{\mathrm{n}2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1\mathrm{n}} & c_{2\mathrm{n}} & \cdots & c_{\mathrm{nn}} \end{bmatrix}$$

Proof:
Consider the determinant of **A** expanded by the first row:

$$\det(A) = a_{11}c_{11} + a_{12}c_{12} + \ldots + a_{1n}c_{1n}$$

Now, we have already seen that if we that if we multiply a row by the cofactors associated with a different row, the result is always zero, e.g. 1^{st} row multiplied by cofactors of second row:

$$a_{11}c_{21} + a_{12}c_{22} + \ldots + a_{1n}c_{2n} = 0$$

So, in general:

$$\sum_{k=1}^{n} a_{ik}c_{jk} = 0 \quad i \neq j$$
$$= |A| \quad i = j$$

Now, if we consider the product:

$$A\ adj(A) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\mathrm{n}} \\ a_{21} & a_{22} & \cdots & a_{2\mathrm{n}} \\ \vdots & \vdots & \ddots & \vdots \\ a_{\mathrm{n}1} & a_{\mathrm{n}2} & \cdots & a_{\mathrm{nn}} \end{bmatrix} \begin{bmatrix} c_{11} & c_{21} & \cdots & c_{\mathrm{n}1} \\ c_{12} & c_{22} & \cdots & c_{\mathrm{n}2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1\mathrm{n}} & c_{2\mathrm{n}} & \cdots & c_{\mathrm{nn}} \end{bmatrix}$$

We can see that we are multiplying *elements* of **A** by *cofactors* of **A**. So using the previous two equations we obtain:

$$A\ adj(A) = \begin{bmatrix} |\mathrm{A}| & 0 & \cdots & 0 \\ 0 & |\mathrm{A}| & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & |\mathrm{A}| \end{bmatrix} = |\mathrm{A}| \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = |\mathrm{A}|\ I$$

Hence:

$$A\ \frac{adj(A)}{|\mathrm{A}|} = I$$

and, therefore:

$$A^{-1} = \frac{adj(A)}{|\mathrm{A}|}$$

This is a simple method of finding the inverse but other methods (e.g. LU decomposition) are usually more efficient.

Example: Find the inverse of:

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

First find det(A), e.g. by expanding down first column:

$$|\mathbf{A}| = \begin{vmatrix} 1 & 3 & -2 \\ 0 & 1 & 6 \\ 1 & -5 & -2 \end{vmatrix} = 1 \times \begin{vmatrix} 1 & 6 \\ -5 & -2 \end{vmatrix} - 0 \begin{vmatrix} 3 & -2 \\ -5 & -2 \end{vmatrix} + 1 \times \begin{vmatrix} 3 & -2 \\ 1 & 6 \end{vmatrix}$$

$$= (1\times(-2) - 6\times(-5)) + ((3\times 6 - (-2)\times 1) = 28 + 20 = 48$$

Then find the matrix of minors, and apply the alternating signs to give the cofactor matrix, and take the transpose to give the adjacent:

$$M = \begin{bmatrix} 28 & -6 & -1 \\ -16 & 0 & -8 \\ 20 & 6 & 1 \end{bmatrix} \quad C = \begin{bmatrix} 28 & 6 & -1 \\ 16 & 0 & 8 \\ 20 & -6 & 1 \end{bmatrix}$$

$$adj(A) = C^T = \begin{bmatrix} 28 & 16 & 20 \\ 6 & 0 & -6 \\ -1 & 8 & 1 \end{bmatrix}$$

so

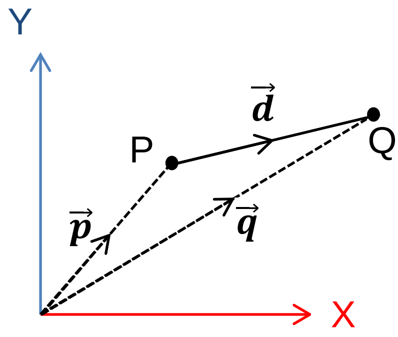
$$\mathbf{A}^{-1} = \frac{adj(\mathbf{A})}{|A|} = \frac{1}{48} \begin{bmatrix} 28 & 16 & 20 \\ 6 & 0 & -6 \\ -1 & 8 & 1 \end{bmatrix}$$

Geometric Representation

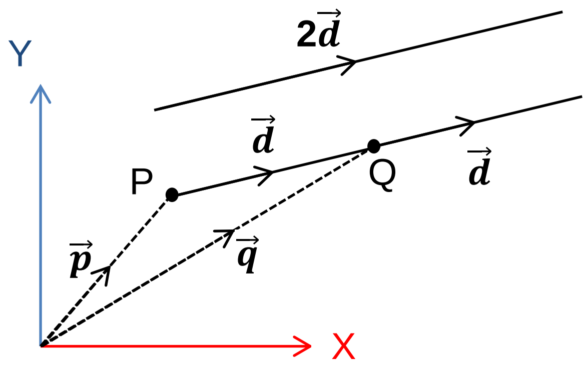
Equation of a Line

As seen previously, a vector can be used to represent a direction in space. For example, if point P is at a coordinate represented by vector **p** and point Q is at a coordinate represented by vector **q**, then the vector **d** representing the displacement from point P to point Q is given by:

$$\vec{d} = \overrightarrow{PQ} = \overrightarrow{PO} + \overrightarrow{OQ} = -\overrightarrow{OP} + \overrightarrow{OQ} = -\vec{p} + \vec{q} = \vec{q} - \vec{p}$$

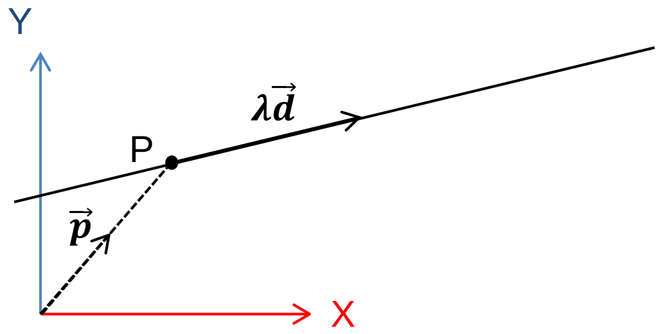


Two times **d** would represent a displacement from P to Q and then the same displacement again:



So by multiplying the vector by a scalar parameter we can find an equation for the infinite set of points which lie on a line parallel to the direction PQ.

Parametric Definition
The specific set of points, **r** say, which lie on a line in the direction PQ and which also pass through points P and Q can be found by first making a fixed displacement from the origin to point P (or Q) and then following this by a scalar multiple of the vector **d** (or –**d**).



$$r = \vec{p} + \lambda\vec{d} = \vec{p} + \lambda(\vec{q} - \vec{p})$$

Every value of the scalar λ gives a point **r** which lies on the line PQ. In particular, in this case, a value of λ=0 will give a point at position P and a value of λ=1 will give a point at position Q. A value of λ=0.5 will give a point half way between P and Q, etc.

Example: Find the parametric equation of a line in the direction **d**=(1,1,1) which passes through point P=(3,-5,2).

$$\text{Equation of line} = \overrightarrow{OP} + \lambda\vec{d} = \begin{pmatrix} 3 \\ -5 \\ 2 \end{pmatrix} + \lambda\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

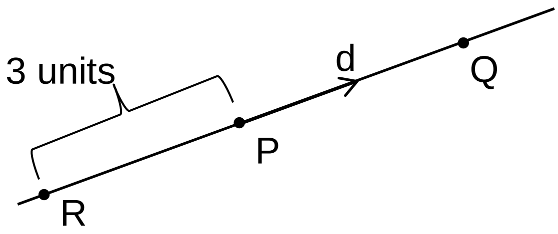
Example: Find the parametric equation of a line which passes through points P=(2,0,1) and Q=(0,2,3).

$$\text{Direction } \vec{d} = \overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \\ 2 \end{pmatrix}$$

$$\text{Equation of line} = \overrightarrow{OP} + \lambda\vec{d} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} + \lambda\begin{pmatrix} -2 \\ 2 \\ 2 \end{pmatrix}$$

Rather than have the parameterisation scaled such that the point moves from P to Q as λ changes from 0 to 1, it is often useful to unitise the direction vector, such that the parameter λ actually refers to the distance from point P towards Q (as opposed to the proportion of the distance from P to Q).

Example: Find the parametric equation of a line which passes through points P=(2,3,2) and Q=(4,4,0) and find the point "R", on the line, which is 3 units further away from Q than from P.



$$\vec{d} = \overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = \begin{pmatrix} 4 \\ 4 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$$

$$|\vec{d}| = \sqrt{2^2 + 1^2 + (-2)^2} = 3 \quad \Longrightarrow \quad \hat{d} = \frac{1}{3}\begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$$

$$\text{Equation of line} = \overrightarrow{OP} + \lambda\hat{d} = \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix} + \lambda\begin{pmatrix} 2/3 \\ 1/3 \\ -2/3 \end{pmatrix}$$

$$R = \overrightarrow{OP} + (-3)\hat{d} = \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix} + (-3)\frac{1}{3}\begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 4 \end{pmatrix}$$

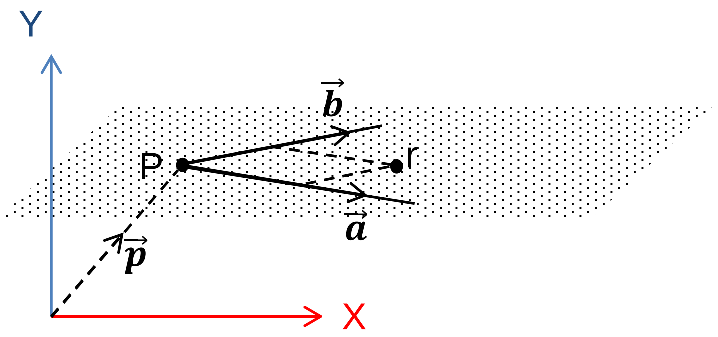
Equation of a Plane

Parametric Definition

Just as a line can be defined by a displacement from the origin to a point on the line, followed by a scalar multiple of a vector in the direction of the line, so a plane can be defined by a displacement from the origin to a point on the plane, followed by a scalar multiple of one vector in the plane, followed by a scalar multiple of another vector in the plane, so long as the two vectors are not parallel.

For example, the point **p** and the two vectors **a** and **b** in the sketch below lie within a plane and **a** and **b** are not parallel. Therefore, any point **r** within the plane can be defined as:

$$\mathbf{r} = \mathbf{p} + \lambda\mathbf{a} + \mu\mathbf{b}$$



For example, we can find the parametric equation of the plane which contains points A=(1,0,1), B=(2,2,2) and C=(5,4,-1).

One line within the plane is

$$\overrightarrow{AB} = \begin{pmatrix}2\\2\\2\end{pmatrix} - \begin{pmatrix}1\\0\\1\end{pmatrix} = \begin{pmatrix}1\\2\\1\end{pmatrix}$$

Another line within the plane is

$$\overrightarrow{AC} = \begin{pmatrix}5\\4\\-1\end{pmatrix} - \begin{pmatrix}1\\0\\1\end{pmatrix} = \begin{pmatrix}4\\4\\-2\end{pmatrix}$$

These lines are not parallel (one is not a multiple of the other, the angle between them according to the dot product is not zero, etc.) so they define a plane.

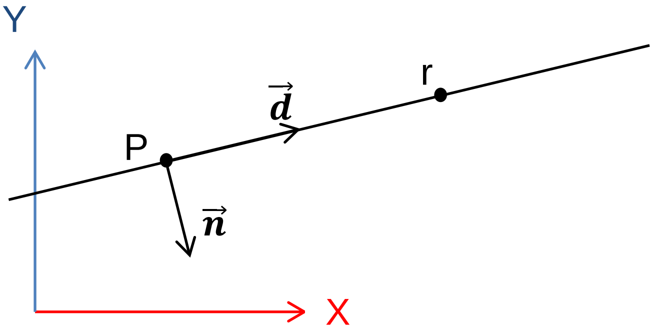
One point on the plane is point A.

Equation of plane is:

$$r = \overrightarrow{OA} + \lambda\,\overrightarrow{AB} + \mu\,\overrightarrow{AC} = \begin{pmatrix}1\\0\\1\end{pmatrix} + \lambda\begin{pmatrix}1\\2\\1\end{pmatrix} + \mu\begin{pmatrix}4\\4\\-2\end{pmatrix}$$

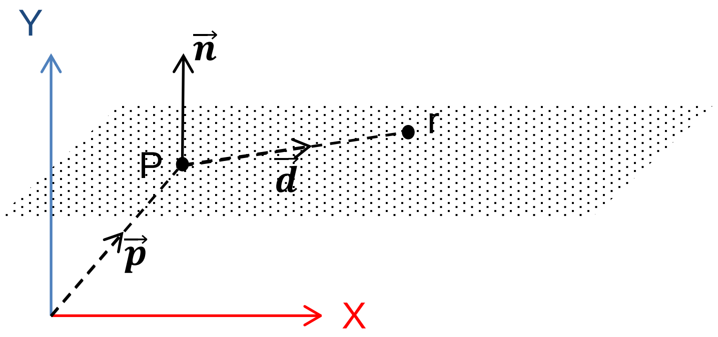
Vector Definition

Another way to define the equation of a plane is to use the dot product. Suppose, in 2D, instead of knowing a point on the line and the direction of a line **d**, a point **p** on the line and the vector **n** normal to the line is known.



For all points **r** along the line, the direction vector **d** could be reconstructed as **d** = **r**–**p**. Since **d** and **n** are perpendicular, the cosine of the angle between them is zero and so the relation **d**·**n**=0 or (**r**-**p**)·**n**=0 can be deduced.

Moving into 3D, in a similar way to a line, a plane can be defined by a point **p** on the plane and a vector **n** normal to that plane. For all points **r** on the plane, once again, the direction vector **d** could be reconstructed as **d** = **r**–**p**.



And again, since the plane can be defined as the set of points perpendicular to **n**, the same relation **d**·**n**=0 or (**r**-**p**)·**n**=0 holds and we can define the equation of a plane normal to **n** and through point **p** to be:

$$(\boldsymbol{r} - \boldsymbol{p}) \cdot \boldsymbol{n} = 0$$

Since the dot product can be multiplied out (it is distributive) then $\boldsymbol{r} \cdot \boldsymbol{n} - \boldsymbol{p} \cdot \boldsymbol{n} = 0$ and the equation of a plane is sometimes written as:

$$\boldsymbol{r} \cdot \boldsymbol{n} = \boldsymbol{p} \cdot \boldsymbol{n}$$

If the vector **r** is represented as **r** = $(x,y,z)^T$ and the dot-product is expanded, the more familiar Cartesian form of the equation of a plane can be found.

Example: Find the vector equation of the plane which passes through point P=(2,0,1) and is normal to the vector n=(3,1,-1). Convert the equation to Cartesian form.

$$\vec{r}\cdot\vec{n} = \vec{p}\cdot\vec{n}$$

$$\vec{r}\cdot\begin{pmatrix}3\\1\\-1\end{pmatrix} = \begin{pmatrix}2\\0\\1\end{pmatrix}\cdot\begin{pmatrix}3\\1\\-1\end{pmatrix} = 6+0-1=5$$

Vector equation:

$$\vec{r}\cdot\begin{pmatrix}3\\1\\-1\end{pmatrix} = 5$$

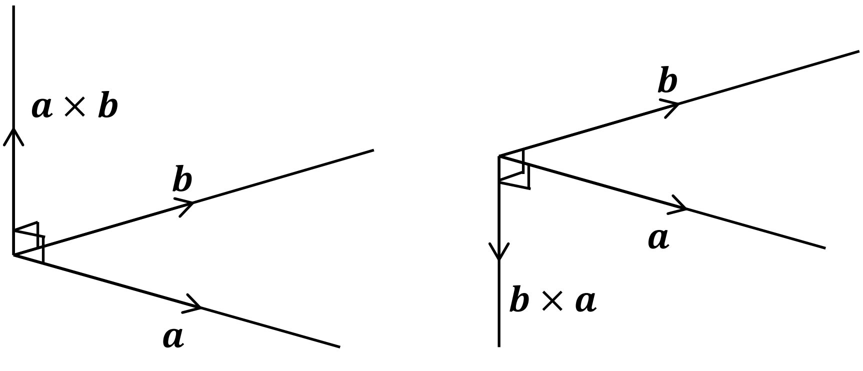
In Cartesian form:

$$\begin{pmatrix}x\\y\\z\end{pmatrix}\cdot\begin{pmatrix}3\\1\\-1\end{pmatrix} = 5 \quad \Longrightarrow \quad 3x+y-z=5$$

Vector (Cross) Product

The vector product is an operator (like + and -) which takes two vectors and turns them into another third vector. It is written as a cross between the two vectors (hence the alternative name *cross product*).

So the vector product of **a** and **b**, say, is $\boldsymbol{a}\times\boldsymbol{b}$ and represents a vector perpendicular to both of the other two and in the direction of a right-handed screw being rotated from vector **a** to **b**.



It follows directly therefore that the cross product is not commutative (the order matters) and in particular:

$$a\times b = -b\times a$$

The cross product can be calculated by evaluating the determinant of the corresponding elements as follows:

$$a \times b = \begin{vmatrix} x & y & z \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} = \begin{pmatrix} a_y b_z - a_z b_y \\ a_z b_x - a_x b_z \\ a_x b_y - a_y b_x \end{pmatrix}$$

And it can be shown that:

$$|a \times b| = |a||b| \sin\theta$$

where θ is the angle between the two vectors.

Which shows that:

$a \times b = 0$ if the two vectors are parallel

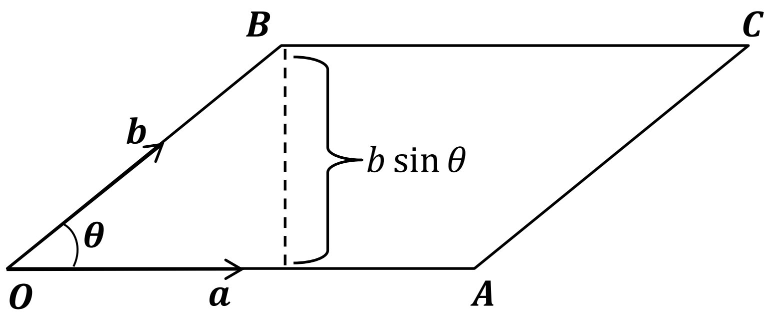
$a \times a = 0$ since vector **a** is parallel to itself, so θ=0

The cross product is *distributive* over addition and subtraction, which means we can multiply out brackets as:

$$(a \pm b) \times c = a \times c \pm b \times c$$

Vector Product as an Area

By considering the parallelogram formed by the two vectors, the modulus of the cross-product can be seen as equal to the area of the parallelogram, or twice the area of the triangle they define.



$$|a \times b| = \text{Area OACB} = 2 \text{ Area OAB}$$

Example: Three coordinates A=(1,1,1), B=(2,0,-2) and C=(4,3,2) define a triangle in 3D space. Find the area of the triangle, and the unit vector perpendicular to it.

Vector of one edge:

$$\overrightarrow{AB} = \mathrm{B} - \mathrm{A} = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -3 \end{pmatrix}$$

Vector of other edge:

$$\overrightarrow{AC} = \mathrm{C} - \mathrm{A} = \begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

Normal to both edges:

$$\vec{n} = \overrightarrow{AB} \times \overrightarrow{AC} = \begin{vmatrix} x & y & z \\ 1 & -1 & -3 \\ 3 & 2 & 1 \end{vmatrix} = \begin{pmatrix} 5 \\ -10 \\ 5 \end{pmatrix}$$

$$|\vec{n}| = \sqrt{5^2 + (-10)^2 + 5^2} = \sqrt{150}$$

$$\text{Area } \Delta\mathrm{ABC} = \frac{1}{2}\,|\vec{n}| = \frac{\sqrt{150}}{2}$$

Perpendicular Unit Vector:

$$\hat{n} = \frac{\vec{n}}{|\vec{n}|} = \frac{1}{\sqrt{150}}\begin{pmatrix} 5 \\ -10 \\ 5 \end{pmatrix} = \frac{1}{5\sqrt{6}}\begin{pmatrix} 5 \\ -10 \\ 5 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}$$

A nice property of the distributive nature of the cross-product makes the calculation of the area of a triangle possible directly from the coordinates, rather than having to derive the edge vectors as above. The area of the triangle with vertices at positions A, B and C is given by:

$$\begin{aligned}
&= \frac{1}{2}\left|\overrightarrow{AB} \times \overrightarrow{AC}\right| \\
&= \frac{1}{2}\left|(\mathrm{B} - \mathrm{A}) \times (\mathrm{C} - \mathrm{A})\right| \\
&= \frac{1}{2}\left|\mathrm{B} \times \mathrm{C} - \mathrm{B} \times \mathrm{A} - \mathrm{A} \times \mathrm{C} + \mathrm{A} \times \mathrm{A}\right| \\
&= \frac{1}{2}\left|\mathrm{B} \times \mathrm{C} + \mathrm{A} \times \mathrm{B} + \mathrm{C} \times \mathrm{A} + 0\mathrm{A}\right| \\
&= \frac{1}{2}\left|\mathrm{A} \times \mathrm{B} + \mathrm{B} \times \mathrm{C} + \mathrm{C} \times \mathrm{A}\right|
\end{aligned}$$

So the area of the triangle defined by three points, A, B and C is given by:

$$Area\ \Delta ABC = \frac{1}{2}|\mathrm{A}\times\mathrm{B}+\mathrm{B}\times\mathrm{C}+\mathrm{C}\times\mathrm{A}|$$

For example, using the three points A, B and C defined in the previous exercise, gives:

$$\mathrm{A}\times\mathrm{B} = \begin{vmatrix} x & y & z \\ 1 & 1 & 1 \\ 2 & 0 & -2 \end{vmatrix} = \begin{pmatrix} -2 \\ 4 \\ -2 \end{pmatrix}$$

$$\mathrm{B}\times\mathrm{C} = \begin{vmatrix} x & y & z \\ 2 & 0 & -2 \\ 4 & 3 & 2 \end{vmatrix} = \begin{pmatrix} 6 \\ -12 \\ 6 \end{pmatrix}$$

$$\mathrm{C}\times\mathrm{A} = \begin{vmatrix} x & y & z \\ 4 & 3 & 2 \\ 1 & 1 & 1 \end{vmatrix} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

$$\mathrm{A}\times\mathrm{B}+\mathrm{B}\times\mathrm{C}+\mathrm{C}\times\mathrm{A} = \begin{pmatrix} -2 \\ 4 \\ -2 \end{pmatrix} + \begin{pmatrix} 6 \\ -12 \\ 6 \end{pmatrix} + \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ -10 \\ 5 \end{pmatrix}$$

$$|\mathrm{A}\times\mathrm{B}+\mathrm{B}\times\mathrm{C}+\mathrm{C}\times\mathrm{A}| = \sqrt{5^2+(-10)^2+5^2} = \sqrt{150}$$

$$\mathrm{Area} = \frac{1}{2}|\mathrm{A}\times\mathrm{B}+\mathrm{B}\times\mathrm{C}+\mathrm{C}\times\mathrm{A}| = \frac{\sqrt{150}}{2}$$

Which thankfully gives the same answer as found using the edge-vectors!

Vector Definition of a Line

The property of a cross-product being zero if the vectors are parallel gives an alternative means of describing the equation of a straight line.

Suppose **r** is a general point on a line which passes through point **p** in direction **d.** Then the parametric form of its equation is:

$$r = \vec{p} + \lambda\vec{d}$$

But we know that (**r–p**) is parallel to **d**, so:

$$(\boldsymbol{r}-\boldsymbol{p})\times\boldsymbol{d} = 0$$

Since the cross product can be multiplied out (it is distributive) then $\boldsymbol{r}\times\boldsymbol{d}-\boldsymbol{p}\times\boldsymbol{d} = 0$ and the equation of a plane is sometimes written as:

$$\boldsymbol{r}\times\boldsymbol{d} = \boldsymbol{p}\times\boldsymbol{d}$$

Example: Find the vector equation of a line which passes through points P=(2,3,2) and Q=(4,4,0).

The direction of the line **d** is:

$$\vec{d} = \overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = \begin{pmatrix} 4 \\ 4 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$$

So the equation of the line can be written in the form $(\boldsymbol{r} - \boldsymbol{p}) \times \boldsymbol{d} = 0$ as:

$$\left(\vec{r} - \begin{pmatrix} 4 \\ 4 \\ 0 \end{pmatrix}\right) \times \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} = 0$$

Or in the form $\boldsymbol{r} \times \boldsymbol{d} = \boldsymbol{p} \times \boldsymbol{d}$ as:

$$\vec{r} \times \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 0 \end{pmatrix} \times \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} = \begin{vmatrix} x & y & z \\ 4 & 4 & 0 \\ 2 & 1 & -2 \end{vmatrix}$$

$$\vec{r} \times \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -8 \\ 8 \\ -2 \end{pmatrix}$$