Matrices

Introduction

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. For example, the internal forces in the members of a pin-jointed truss can be directly related to the externally applied loading via a series of simultaneous equations. These equations can be represented in matrix form, and matrix methods can be used to solve for the forces. Matrices are particularly useful for analysing complex linear structures by computational or numerical methods.

Outline:

- Determinants
  - Second Order
  - Third Order
  - Properties of Determinants
  - Higher Order Determinants
- Matrices
  - Matrix Algebra
  - Matrix Multiplication
  - Representation of linear equations
  - Transpose
  - Special Types of Matrices
- Inverse of a Matrix
  - Adjoint matrix
- Solution of Linear Equations
  - Gaussian Elimination
  - Gaussian Elimination with Pivoting
- Uniqueness and Consistency
  - Homogenous Equations
- Eigenvectors and Eigenvalues
  - Properties of Eigenvalues
  - Engineering Application

Determinants

Second Order

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:

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

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:

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

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":

$$\left| \begin{matrix} a_{11} & a_{12} & a_{13} & (a_{11}) & (a_{12}) \\ a_{21} & a_{22} & a_{23} & (a_{21}) & (a_{22}) \\ a_{31} & a_{32} & a_{33} & (a_{31}) & (a_{32}) \end{matrix} \right|$$

i.e.:

$$\begin{aligned} D = &+a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &- a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \\ = &+a_{11}(a_{22}a_{33} - a_{23}a_{32}) \\ &- a_{12}(a_{21}a_{33} - a_{23}a_{31}) \\ &+ a_{13}(a_{21}a_{32} - a_{22}a_{31}) \end{aligned}$$

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

$$\begin{aligned} D_3 &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13} \end{aligned}$$

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.:

$$M_{11} = \left| \begin{array}{ccc} a_{11} & \times & \times \\ \times & a_{22} & a_{23} \\ \times & a_{32} & a_{33} \end{array} \right|$$

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:

$$M_{21} = \left| \begin{array}{ccc} \times & a_{12} & a_{13} \\ a_{21} & \times & \times \\ \times & a_{32} & a_{33} \end{array} \right|$$

So, in this case the determinant is:

$$\begin{aligned} D_3 &= -a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{22}\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{23}\begin{vmatrix} a_{11} & a_{32} \\ a_{31} & a_{32} \end{vmatrix} \\ &= -a_{21}M_{21} + a_{22}M_{22} - a_{23}M_{23} \end{aligned}$$

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 determinant and form 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)

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

= 4x(0-4) + 2x(-6+1) + 3x(12-0)
= -16 - 10 + 36
= 10

b)

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

= 3x(12-0) + 1x(-16+2) + 2x(0-6)
= 36 - 14 - 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:

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

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:

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

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

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

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.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \end{vmatrix} = 0$$

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.

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

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 determinant 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 3^{rd} order determinant:

$$D = \begin{vmatrix} a_{11} & a_{12} & \ldots\ldots & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots & a_{2n} \\ \vdots & & & \\ \vdots & & & \\ a_{n1} & a_{n2} & \ldots\ldots & a_{nn} \end{vmatrix} = \sum_{j=1}^{n} a_{ij}c_{ij}$$

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*.

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

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

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \left|\begin{array}{c|cc|} +a_{11} & \times & \times \\ \times & a_{22} & a_{23} \\ \times & a_{32} & a_{33} \end{array}\right| + \left|\begin{array}{|ccc|} \times & -a_{12} & \times \\ a_{21} & \times & a_{23} \\ a_{31} & \times & a_{33} \end{array}\right| + \left|\begin{array}{|cc|c} \times & \times & +a_{13} \\ a_{21} & a_{22} & \times \\ a_{31} & a_{32} & \times \end{array}\right|$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = \left|\begin{array}{c|ccc|} +a_{11} & \times & \times & \times \\ \times & a_{22} & a_{23} & a_{24} \\ \times & a_{32} & a_{33} & a_{34} \\ \times & a_{42} & a_{43} & a_{44} \end{array}\right| + \left|\begin{array}{|cccc|} \times & -a_{12} & \times & \times \\ a_{21} & \times & a_{23} & a_{24} \\ a_{31} & \times & a_{33} & a_{34} \\ a_{41} & \times & a_{43} & a_{44} \end{array}\right| + \ldots$$

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:

$$\begin{vmatrix} 2 & 3 & 5 & -1 \\ 1 & 2 & -3 & 2 \\ 4 & -1 & 2 & 5 \\ 2 & -2 & 3 & 1 \end{vmatrix} = \begin{vmatrix} 0 & -1 & 11 & -5 \\ 1 & 2 & -3 & 2 \\ 0 & -9 & 14 & -3 \\ 0 & -6 & 9 & -3 \end{vmatrix} \begin{matrix} \Leftarrow -2xRow2 \\ \\ \Leftarrow -4xRow2 \\ \Leftarrow -2xRow2 \end{matrix}$$

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:

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

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

$$= -\begin{vmatrix} -1 & 11 & -5 \\ -9 & 14 & -3 \\ -6 & 9 & -3 \end{vmatrix} = -\begin{vmatrix} -1 & 11 & -5 \\ 0 & 85 & 42 \\ 0 & -57 & 27 \end{vmatrix} \begin{matrix} \\ \Leftarrow -9xRow1 \\ \Leftarrow -6xRow1 \end{matrix}$$

Expanding down first column therefore gives:

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

= -85x27 – 42x(-52)
= 99

Matrices

A matrix is a table of numbers arranged in a bracketed rectangular array of rows and columns. For example, a matrix **A** might be:

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

Matrices are usually denoted by bold capitals (in typescript) and by capitals with a wavy line underneath (in handwriting):

$$\mathbf{A} = \underset{\sim}{A} \text{ or } \underline{A} \text{ or } [A]$$

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 only if they have the same order and if all the corresponding elements are equal.

Unlike determinants, matrices have no algebraic value. They should be regarded as an *operator* in the same way as *d/dx*, or *i*.

Matrix Algebra

Before showing how we use matrices and how they relate with determinants, we will first look at elementary algebra. Algebraic manipulation of matrices is known as *linear algebra*.

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:

$$\begin{bmatrix} 5 & 3 & -2 \\ 1 & 4 & 2 \end{bmatrix} + \begin{bmatrix} 4 & 0 & -9 \\ -1 & 8 & 5 \end{bmatrix} = \begin{bmatrix} 9 & 3 & -11 \\ 0 & 12 & 7 \end{bmatrix}$$

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:

$$\begin{bmatrix} 5 & 3 & -2 \\ 1 & 4 & 2 \end{bmatrix} - \begin{bmatrix} 4 & 0 & -9 \\ -1 & 8 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 7 \\ 2 & -4 & -3 \end{bmatrix}$$

Multiplication by a scalar

Multiplying by a scalar simply multiplies *each element* of the matrix by the scalar:

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

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 multiplying the first *row* of **A** by the second *column* of **B**, and so on. e.g.:

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

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

More formally, if **AB = C** 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:

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

Representation of linear equations

Matrix multiplication is useful for representing systems of linear simultaneous equations. For example:

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

can be written:

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

Transpose

The transpose of a matrix **A** is denoted by $\mathbf{A}^T$ (or sometimes **A**') and is formed by *swapping rows* and *columns.* So, if:

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

Special Types of Matrices

Row and column vectors

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

$$\underline{\mathbf{a}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \quad \text{and} \quad \underline{\mathbf{a}}^T = \begin{bmatrix} a_1 & a_2 & \ldots & a_m \end{bmatrix}$$

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), $\mathbf{O}_{mn}$, e.g.:

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

Clearly, adding a matrix **A** to the appropriate order zero matrix leaves **A** 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. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

The identity matrix

The *identity matrix* (or unit matrix) is very important in matrix algebra. It is a *square* matrix of order (m×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 with a (m×m) matrix **A**, there is no change to **A**:

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

More generally, if **A** is an (m×n) matrix, then:

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

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.

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 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:

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

Now, recalling the lecture on *determinants*, the quantity *ad - bc* is the *determinant* of the (*2*×*2*) matrix **A**, written **|A|** or *det*(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:

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

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

$$|A| = ad - bc = 3\times2 - 7\times(-5) = 6 + 35 = 41$$

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

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

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

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**:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots\dots & a_{1m} \\ a_{21} & a_{22} & \dots\dots & a_{2m} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots\dots & a_{mm} \end{bmatrix}$$

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

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

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \dots\dots & c_{1m} \\ c_{21} & c_{22} & \dots\dots & c_{2m} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ c_{m1} & c_{m2} & \dots\dots & c_{mm} \end{bmatrix}$$

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

$$adj(A) = \mathbf{C}^T = \begin{bmatrix} c_{11} & c_{21} & \dots\dots & c_{m1} \\ c_{12} & c_{22} & \dots\dots & c_{m2} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ c_{1m} & c_{2m} & \dots\dots & c_{mm} \end{bmatrix}$$

Proof

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

$$\det(A) = a_{11}c_{11} + a_{12}c_{12} + \dots + {}_{a1}nc_{1n}$$

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. 1st row multiplied by cofactors of second row:

$$a_{11}c_{21} + a_{12}c_{22} + \dots + {}_{a1}nc_{2n} = 0$$

So, in general:

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

Now, if we consider the product:

$$\mathbf{A}\,adj(\mathbf{A}) = \begin{bmatrix} a_{11} & a_{12} & ..... & a_{1m} \\ a_{21} & a_{22} & ..... & a_{2m} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & ..... & a_{mm} \end{bmatrix} \begin{bmatrix} c_{11} & c_{21} & ..... & c_{m1} \\ c_{12} & c_{22} & ..... & c_{m2} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ c_{1m} & c_{2m} & ..... & c_{mm} \end{bmatrix}$$

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

$$\mathbf{A}\,adj(\mathbf{A}) = \begin{bmatrix} |A| & 0 & ..... & 0 \\ 0 & |A| & ..... & 0 \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ 0 & \ldots & ..... & |A| \end{bmatrix} = |A| \begin{bmatrix} 1 & 0 & ..... & 0 \\ 0 & 1 & ..... & 0 \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ 0 & \ldots & ..... & 1 \end{bmatrix}$$

Hence:

$$\mathbf{A}\,adj(\mathbf{A}) = |A|\,I$$

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

and, therefore:

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

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:

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

= (1x(-2) – 6x(-5)) + ((3x6 – (-2)x1) = 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

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

Solution of Linear Equations

As we have seen, a set of linear equations:

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2$$
$$a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n$$

Can be written in matrix form as:

$$\begin{bmatrix} a_{11} & a_{12} & \ldots\ldots & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots & a_{2n} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \ldots\ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_n \end{bmatrix}$$

or:

$$\mathbf{A}\,\underline{x} = \underline{b}$$

We know the values of **A** and **b** and we want to find **x**. The following procedure *could* be used:

1. Find the inverse of **A**, i.e. $\mathbf{A}^{-1}$.
2. Pre multiply both sides of **Ax = b** by $\mathbf{A}^{-1}$ (remember that the order of multiplication is important!).

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

$$\mathbf{I}\,\underline{x} = \mathbf{A}^{-1}\,\underline{b}$$

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

This seems fine, but in practice this method is seldom used since there is a lot of tedious work involved in finding the inverse matrix (and numerical problems can arise). However, the principle is important, as we shall see later on. The following methods for solving sets of linear equations are more practical.

Gaussian Elimination

Consider the usual method of solving simultaneous equations, for example:

$$\begin{aligned} 2x + 3y + 4z &= 23 && \text{(Row1)} \\ 4x + 10y + 9z &= 59 && \text{(Row2)} \\ 6x + 17y + 20z &= 101 && \text{(Row3)} \end{aligned}$$

We first eliminate the *x* terms from the 2nd and 3rd equations:

$$\begin{aligned} 2x + 3y + 4z &= 23 && \text{(Row1)} \\ 4y + z &= 13 && \text{(Row2 – 2xRow1)} \\ 8y + 8z &= 32 && \text{(Row3 – 3xRow1)} \end{aligned}$$

Now we eliminate the *y* term from the 3rd equation:

$$\begin{aligned} 2x + 3y + 4z &= 23 \\ 4y + z &= 13 \\ 6z &= 6 && \text{(Row3 – 2xRow2)} \end{aligned}$$

and finally *back substitute* to obtain:

z = 1
y = 3
x = 5

The equivalent method in matrix notation is called *Gaussian Elimination*. For example, we can write the previous example in matrix form as:

$$\begin{bmatrix} 2 & 3 & 4 \\ 4 & 10 & 9 \\ 6 & 17 & 20 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 23 \\ 59 \\ 101 \end{bmatrix}$$

We now write a Gaussian elimination *Tableau*, which consists of the matrix augmented by the right hand side column vector:

$$\left[\begin{array}{ccc|c} 2 & 3 & 4 & 23 \\ 4 & 10 & 9 & 59 \\ 6 & 17 & 20 & 101 \end{array}\right]$$

Next we perform row operations to make the bottom left of the tableau equal to zero, e.g. first get zeros in the first column below the diagonal:

$$\left[\begin{array}{ccc|c} 2 & 3 & 4 & 23 \\ 0 & 4 & 1 & 13 \\ 0 & 8 & 8 & 32 \end{array}\right] \begin{array}{l} \\ \Leftarrow Row2 - 2Row1 \\ \Leftarrow Row3 - 3Row1 \end{array}$$

Then get zeros in the second column below the diagonal:

$$\left[\begin{array}{ccc|c} 2 & 3 & 4 & 23 \\ 0 & 4 & 1 & 13 \\ 0 & 0 & 8 & 32 \end{array}\right] \begin{array}{l} \\ \\ \Leftarrow Row3 - 2Row2 \end{array}$$

So, our original matrix equation after the Gaussian elimination process can be written as:

$$\begin{bmatrix} 2 & 3 & 4 \\ 0 & 4 & 1 \\ 0 & 0 & 8 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 23 \\ 13 \\ 32 \end{bmatrix}$$

Finally, back substitute to obtain *x*, *y* and *z*, i.e. starting with the last row:

$6z = 6 \implies z = 1$
$4y + 1z = 4y + 1x1 = 13 \implies 4y = 12 \implies y = 3$
$2x + 3y + 4z = 2x + 9 + 4 = 23 \implies 2x = 10 \implies x = 5$

therefore:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix}$$

So, to summarise:

1. Write the tableau by augmenting the matrix with the right hand side of the equation.
2. Subtract suitable multiples of row 1 from each of the rows below in order to make the elements in column 1 (below row 1) equal to zero.

3. Do the same with row/column two, i.e. Subtract suitable multiples of row 2 from each of the rows below in order to make the elements in column 2 (below row 2) equal to zero.
4. Continue this process until the bottom left triangle of the tableau consists of zeros.
5. Back substitute to find the unknowns, starting with the bottom row.
6. Gaussian elimination can fail if, when subtracting from the i^{th} row to produce zero in the $(i-1)^{th}$ column, the element in the i^{th} column also becomes zero. There can also be problems of numerical accuracy when performing Gaussian elimination computationally if the i^{th} element is very small. A technique for overcoming these problems is called *partial pivoting*.

Gaussian Elimination with Pivoting

Consider the system of equations:

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

Writing in augmented tableau form we get:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 7 \\ 2 & -1 & 2 & 4 \\ 3 & -2 & 4 & 7 \end{array}\right]$$

With partial pivoting we examine the *first column* of the tableau and find the element with the largest modulus (i.e. largest value independent of sign). If this is *not* in the first row then we interchange the first row with the row containing the largest element in the first column. This largest number then appears on the diagonal and is called the pivot since the matrix changes or 'pivots' about this number. We then eliminate as usual.

Next we look at the second column and to find the largest modulus below the first row. Again, we swap the row with the largest value in the second column with the second row and then eliminate as usual.

This proceeds until the full elimination procedure is complete. We can then back substitute to find the unknowns.

Example:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 7 \\ 2 & -1 & 2 & 4 \\ 3 & -2 & 4 & 7 \end{array}\right]$$

Here, the largest value in the first column is in the *third* row (3), so we swap this with the first row:

$$\left[\begin{array}{ccc|c} 3 & -2 & 4 & 7 \\ 2 & -1 & 2 & 4 \\ 1 & 2 & 1 & 7 \end{array}\right]$$

Now we can perform row operations to eliminate the elements of the first column below the first row:

$$\left[\begin{array}{ccc|c} \mathbf{3} & -2 & 4 & 7 \\ 0 & 1/3 & -2/3 & -2/3 \\ 0 & 8/3 & -1/3 & 14/3 \end{array}\right] \begin{array}{l} \\ \Leftarrow Row2 - 2/3\, Row1 \\ \Leftarrow Row3 - 1/3\, Row1 \end{array}$$

Now, the largest value in the second column (below the first row) is in the third row, so we swap this with the second row:

$$\left[\begin{array}{ccc|c} 3 & -2 & 4 & 7 \\ 0 & \mathbf{8/3} & -1/3 & 14/3 \\ 0 & 1/3 & -2/3 & -2/3 \end{array}\right]$$

And perform row operations to eliminate the elements in the second column below the second row:

$$\left[\begin{array}{ccc|c} 3 & -2 & 4 & 7 \\ 0 & \mathbf{8/3} & -1/3 & 14/3 \\ 0 & 0 & -5/8 & -5/4 \end{array}\right] \quad \Leftarrow Row3 - \tfrac{1}{8}\,Row2$$

The elimination is now complete (although we would carry on the procedure if the matrix was larger). We now back substitute to find the unknown values of **x**:

-5/8 z = -5/4 $\Longrightarrow$ z = 2

8/3 y – 1/3 z = 8/3 y – 2/3 = 14/3 $\Longrightarrow$ 8/3 y = 16/3 $\Longrightarrow$ y = 2

3x - 2y + 4z = 3x - 4 + 8 = 7 $\Longrightarrow$ 3x = 3 $\Longrightarrow$ x = 1

Gaussian elimination is much more efficient than using matrix inversion, but we need to keep track of the row operations and hence information about the right hand side. Therefore, if we wish to solve the equation with different right hand sides we need to resolve a set of equations each time. A more suitable method in this case is to use *LU decomposition*.

Uniqueness and Consistency

We have seen that the solution of the equation **A**$\underline{x}$ **=** $\underline{b}$ could be solved using matrix inversion, i.e. calculating:

$$x = \mathbf{A}^{-1}b = \frac{adj(\mathbf{A})}{|A|}b$$

What happens to the solution of **Ax=b** when |**A**| = 0? There are two possibilities:

1. One (or more) equation(s) of the set represented by the matrix equation is a linear combination of the others (i.e. there are only *n-1* (or fewer) independent equations). In this case there are an *infinite* number of solutions for **x**. For example, consider the following equations:

   2x - y + 3z = 4
   x + 3y - z = 5
   3x + 2y + 2z = 9

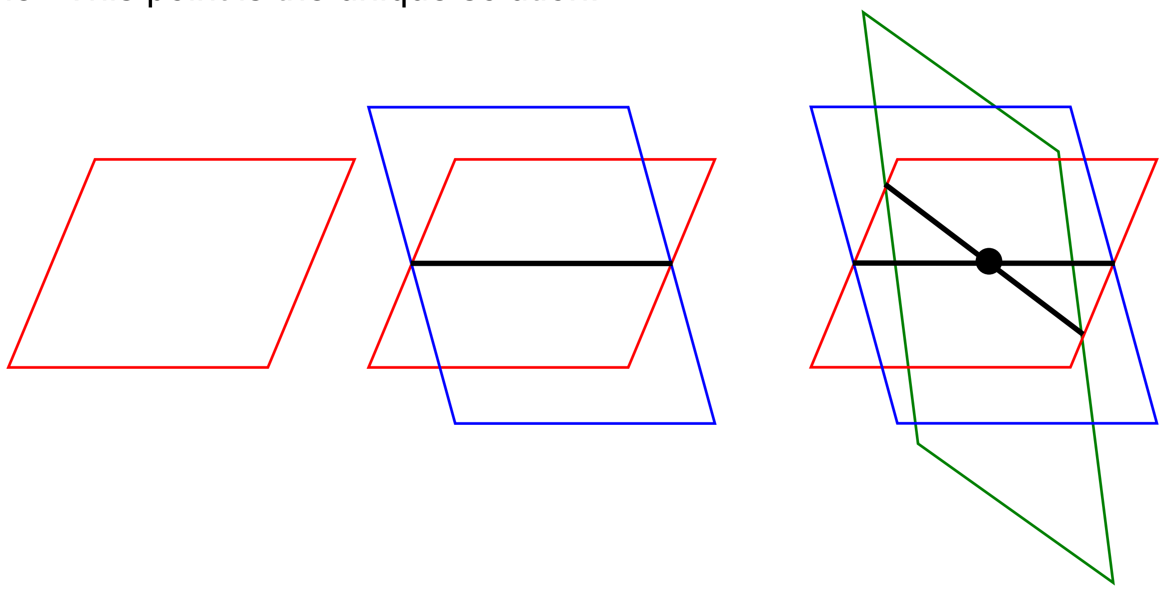
   Here $R_3 = R_1 + R_2$ and hence |**A**| = 0.

2. The equations are *inconsistent* and there is no solution. For example, consider:

$$\begin{aligned} 2x - y + 3z &= 4 \\ x + 3y - z &= 5 \\ 3x + 2y + 2z &= \mathbf{8} \end{aligned}$$

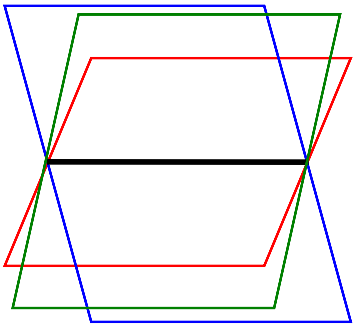
   Again |**A**| = 0 since the third equation is inconsistent with the sum of the other two (i.e. there are no values of *x*, *y* and *z* which satisfy the first two equations which also satisfy the third equation).

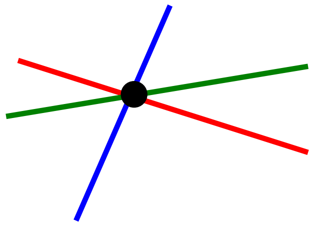
To help us understand this physically, we can represent these possibilities geometrically for 3×3 matrices. Each equation represents a *plane in three-dimensional space*. If |**A**| ≠ 0 then a unique solution exists. The first two planes intersect along a line and the third plane intersects at a point along the line. This point is the unique solution:



In the situation where |**A**| = 0 we either have:

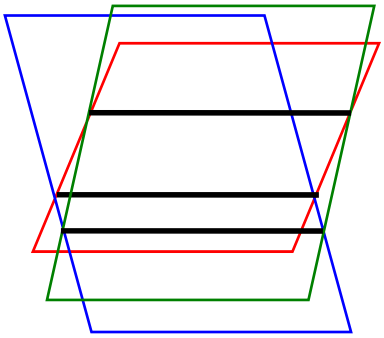
1. All three planes intersecting on the same line. Thus, there are an *infinite* number of possible solutions lying along the line:

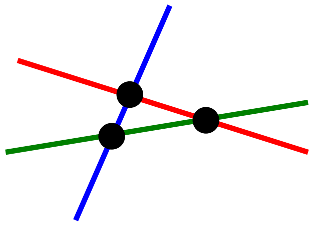




Or:

2. The third plane lies *parallel* to the line of intersection of the other two planes. Thus, no point lies on *all three* planes:





Homogenous Equations

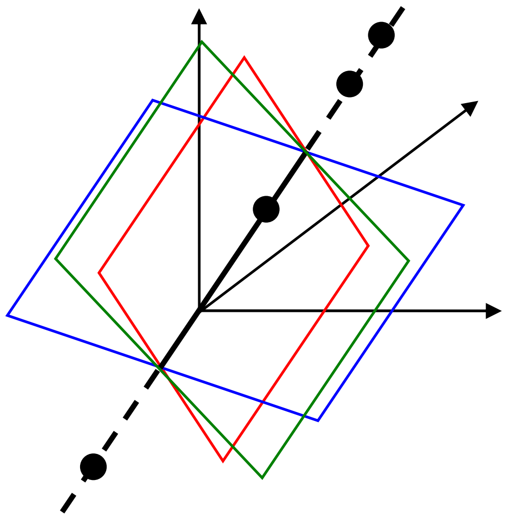
For the special case where **b** = **0**, the equation is known as homogenous:

$$\mathbf{A}\,\underline{x} = 0$$

Two possible solutions may exist:

If |**A**| ≠ 0 then there exists only the trivial solution **x = 0**.

If |**A**| = 0 then the possibility of inconsistent equations *cannot* exist (the right-hand-side is always consistent with linear combinations of the equations). Hence we must have an *infinite* number of solutions. In this case the line of the infinite solutions must pass through the point (0, 0, 0) (since the zero vector **x = 0** is clearly a solution). So, the solution must have components of fixed ratio. (i.e. the solution will be the equation of a line passing through zero. We can multiply the *x*, *y* and *z* values for one particular solution by a common factor, and this will be another solution).



Eigenvectors and Eigenvalues

The basic problem is to find values of scalars λ_i and corresponding column vectors, $\mathbf{x}_i$ which satisfy the equation:

$$\mathbf{A}\,\underline{x} = \lambda\,\underline{x}$$

where **A** is a square matrix. Thus, multiplying $\mathbf{x}_i$ by **A** has the same effect as multiplying $\mathbf{x}_i$ by the scalar λ_i. The significance of this will become clear.

We can rewrite the equation as: $\mathbf{A}\,\underline{x} = \lambda\, I\,\underline{x}$ since multiplying by the identity matrix **I** has no effect. Rearranging, we can write:

$$(\mathbf{A} - \lambda\, I)\underline{x} = \underline{0}$$

This equation is *homogenous*, i.e. the right hand side is a zero column vector, **0**, and the left hand side is an unknown column vector, **x**, multiplied by a square matrix, (**A**-λ**I**). We have seen that a homogenous equation of this type has non-trivial solutions *only* if its determinant is zero:

$$|\mathbf{A} - \lambda\, I| = 0$$

This equation is known as the *characteristic equation* of the matrix **A**. For an (n×n) matrix the characteristic equation is an n^{th} order polynomial in λ.

If we calculate the *roots* of this equation we find *n* values of λ. These values of λ are called *eigenvalues* of **A**. Each eigenvalue, λ_i, produces a corresponding solution $\mathbf{x}_i$. The vector, $\mathbf{x}_i$, is called an *eigenvector* of **A**. These eigenvectors may be found by substituting the corresponding eigenvalue into the original equation. This is best explained through an example.

Example: Find the eigenvalues and eigenvectors of matrix **A** = $\begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 0 \\ 1 & 0 & 3 \end{bmatrix}$

The characteristic equation is:

$$|\mathbf{A} - \lambda I| = \begin{vmatrix} 3-\lambda & 2 & 1 \\ 2 & 3-\lambda & 0 \\ 1 & 0 & 3-\lambda \end{vmatrix} = 0$$

Expanding along the bottom row gives:

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

$$= 1 \times [0 - (3-\lambda)] - 0[\quad] + (3-\lambda) \times [(3-\lambda)(3-\lambda) - 4]$$

$$= -(3-\lambda) + (3-\lambda)[(3-\lambda)(3-\lambda) - 4]$$

$$= (3-\lambda)\left[(3-\lambda)^2 - 5\right]$$

$$= (3-\lambda)(9 - 6\lambda + \lambda^2 - 5)$$

$$= (3-\lambda)(\lambda^2 - 6\lambda + 4) \qquad = 0$$

$\therefore \lambda = 3, 3+\sqrt{5}, 3-\sqrt{5}$ are our 3 eigenvalues of **A**

Now we can find the three corresponding eigenvalues:

Firstly take λ=3:

$$A-\lambda I = A-3I$$

$$(A-3I)x = 0$$

$$\begin{bmatrix} 0 & 2 & 1 \\ 2 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Notice that Row2 is 2xRow3. Hence the rows are not independent.

Solving for the middle row gives 2 $x_1 = 0 \Longrightarrow x_1 = 0$ (also given by solving for the bottom row).

Solving for the top row gives 2 $x_2 + x_3 = 0 \Longrightarrow x_3 = -2\, x_2$

So we have a ratio between the values of x. The eigenvector is:

$$x = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix} \mu$$

where μ is any real number, i.e. any multiple of the vector is a valid eigenvector.

The eigenvectors corresponding to the other eigenvalues are:

$$\lambda = 3+\sqrt{5} \quad \Rightarrow \quad x = \begin{bmatrix} \sqrt{5} \\ 2 \\ 1 \end{bmatrix} \mu \qquad \lambda = 3-\sqrt{5} \quad \Rightarrow \quad x = \begin{bmatrix} \sqrt{5} \\ -2 \\ -1 \end{bmatrix} \mu$$

Properties of Eigenvalues

Two useful properties can be derived allowing a simple check that the correct eigenvalues have been found:

The *sum* of the eigenvalues is equal to the *trace* of the matrix **A**, where the trace is defined as the sum of the elements along the leading diagonal of **A**.

$$\lambda_1+\lambda_2+\ldots+\lambda_n = a_{11}+a_{22}+\ldots+a_{nn} = \text{trace}(A)$$
$$\sum_{i=1}^{n}\lambda_i = \sum_{i=1}^{n} a_{ii}$$

The *product* of the eigenvalues is equal to the determinant of the matrix **A**.

$$\lambda_1\,\lambda_2\,\ldots\,\lambda_n = |\mathbf{A}| = \det(A)$$
$$\prod_{i-1}^{n}\lambda_i = |A|$$

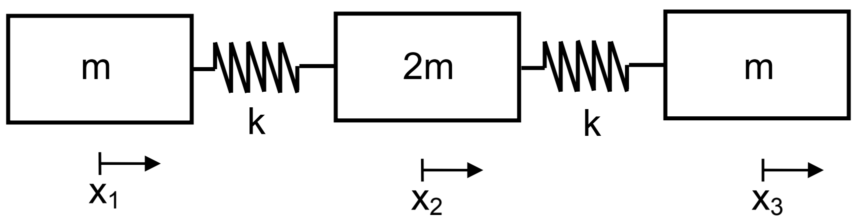
Example: Check the eigenvalues from the previous example.

$$\lambda_1+\lambda_2+\lambda_3 = 3+(3+\sqrt{5})+(3-\sqrt{5}) = 9$$
$$a_{11}+a_{22}+a_{33} = 3+3+3 = 9$$

$$\lambda_1\,\lambda_2\,\lambda_3 = 3(3+\sqrt{5})(3-\sqrt{5}) = 3(9 - 3\sqrt{5} + 3\sqrt{5} - 5) = 3x4 = 12$$
$|\mathbf{A}-\lambda\mathbf{I}| = (3-\lambda)(\lambda^2-6\lambda+4)$ as calculated in example above
so $\lambda=0 \Longrightarrow |A| = (3-0)(0-0+4) = 3x4 = 12$

Engineering Application

Consider the mechanical spring/mass system shown below:



Many structural systems can be simplified and represented in this way.

The mass of each particle is *m* and 2*m*, and *k* is the stiffness of each connecting spring. The system is said to have 3 *degrees of freedom*, since three co-ordinates (x_1, x_2, x_3) are required to describe its displacement. By application of Newton's second law to each mass in turn, we can obtain equations of motion:

$$m\ddot{x}_1 = -kx_1 + kx_2$$
$$2m\ddot{x}_2 = kx_1 - 2kx_2 + kx_3$$
$$m\ddot{x}_3 = kx_2 - kx_3$$

In the case where there are no external forces, these equations can be written in matrix form as:

$$\begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \end{bmatrix} = -\begin{bmatrix} k/m & -k/m & 0 \\ -k/2m & k/m & -k/2m \\ 0 & -k/m & k/m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

So, in general we can write the equations of motion of an *n* degree of freedom system as:

$$\underline{\ddot{x}} = -A\underline{x}$$

where **A** is an ($n \times n$) matrix. The system will vibrate at a number of *natural frequencies*, where, for a given natural frequency each mass moves at the same frequency but with differing amplitudes. The ratio of these amplitudes is known as the *mode shape*. We can write each component of **x** as:

$$x_i(t) = x_i = X_i \sin(\omega t)$$

where ω is the natural frequency of vibration and X_i is the amplitude of the motion described by x_i. If we differentiate this twice with respect to time, *t*, to give acceleration, we get:

$$\ddot{x} = -\omega^2 X_i \sin(\omega t)$$

So, substituting into the equation of motion we get:

$$-\omega^2 \underline{X} \sin(\omega t) = -A\underline{x} = -A\underline{X} \sin(\omega t)$$

or

$$A\underline{X} = \omega^2 \underline{X}$$ where **X** is a column vector of X_i values.

This corresponds to our eigenvalue problem with $\lambda = \omega^2$! Thus, for non-trivial solutions, ω^2 must be the eigenvalues of **A** and **X** must be the corresponding eigenvectors. Thus we can find the natural frequencies, ω, and mode shapes.

So, for our example, we require:

$$\left|\mathbf{A}-\omega^2\mathbf{I}\right| = 0$$

$$\begin{vmatrix} k/m-\omega^2 & -k/m & 0 \\ -k/2m & k/m-\omega^2 & -k/2m \\ 0 & -k/m & k/m-\omega^2 \end{vmatrix} = 0$$

If we expand along the first row:

$$\left(\frac{k}{m}-\omega^2\right)\left[\left(\frac{k}{m}-\omega^2\right)^2-\left(\frac{k^2}{2m^2}\right)\right]-\left(\frac{-k}{m}\right)\left[\frac{-k}{2m}\left(\frac{k}{m}-\omega^2\right)\right]+0\left[\quad\right]=0$$

$$\left(\frac{k}{m}-\omega^2\right)\left[\frac{k^2}{m^2}-\frac{2k\omega^2}{m}+\omega^4-\frac{k^2}{2m^2}\right]-\left(\frac{k^2}{2m^2}\right)\left[\frac{k}{m}-\omega^2\right]=0$$

$$\cancel{\frac{k^3}{m^3}}-\frac{2k^2\omega^2}{m^2}+\frac{k\omega^4}{m}-\cancel{\frac{k^3}{2m^3}}-\cancel{\frac{k^2\omega^2}{m^2}}+\frac{2k\omega^4}{m}-\omega^6+\cancel{\frac{k^2\omega^2}{2m^2}}-\cancel{\frac{k^3}{2m^3}}+\cancel{\frac{k^2\omega^2}{2m^2}}=0$$

$$-\omega^6+\frac{3k}{m}\omega^4-\frac{2k^2}{m2}\omega^2=0$$

Factoring out gives:

$$\omega^2\left(\omega^2-\frac{k}{m}\right)\left(\omega^2-\frac{2k}{m}\right)=0$$

$$\Rightarrow\quad \omega^2=0,\ \frac{k}{m},\ \frac{2k}{m}$$

$$\Rightarrow\quad \omega=0,\ \sqrt{\frac{k}{m}},\ \sqrt{\frac{2k}{m}}$$

The corresponding eigenvectors are:

$$\omega = 0 \Rightarrow \begin{bmatrix} k/m & -k/m & 0 \\ -k/2m & k/m & -k/2m \\ 0 & -k/m & k/m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Row1 $\Longrightarrow x_1 = x_2$

Row3 $\Longrightarrow x_2 = x_3$

$$\Rightarrow \quad X_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \mu$$

Similarly it can be shown that $X_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \mu$ and $X_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \mu$

Physically these represent modes of vibration:

ω_1=0

ω_2=√(k/m)

ω_3=√(2k/m)