Why are they useful?

We often wish to represent a system of linear equations (equations which don’t involve sin, ln, squares etc.) in a compact form. Matrix algebra is one way of doing so. In particular, it allows us to solve large numbers of simultaneous equations by so-called Gaussian elimination (not covered here).

Matrices also appear in structural geology, where we are concerned with deforming objects, and also in some numerical solutions (implicit schemes) of differential equations. Matrix algebra also forms the basis of tensor analysis, which is used in deriving e.g. the fundamental equations of elasticity and relativity.

Example Here’s an example of a set of linear equations.

A rock sample contains three oxides, measured in the lab:

SiO$_2$ 11 parts
MgO 17 parts
CaO 1 part

We want to establish what proportions the following minerals are present in:

Forsterite $\text{Mg}_2\text{SiO}_4$ = 2 MgO + $\text{SiO}_2$
Enstatite $\text{MgSiO}_3$ = MgO + $\text{SiO}_2$
Diopside $\text{CaMgSi}_2\text{O}_6$ = CaO + MgO + 2$\text{SiO}_2$

Let the amounts of forsterite, enstatite and diopside by $f$, $e$ and $d$ respectively. Then we have

$\text{MgO} : 2f + e + d = 17$
$\text{SiO}_2 : f + e + 2d = 11$
$\text{CaO} : d = 1$

So we have three equations and three unknowns. How would we solve this?
Let’s return to the question of solving linear simultaneous equations. For instance, let’s say we have three (coupled) equations in three unknowns \(x, y\) and \(z\) which we wish to solve. Rather than writing the equations

\[
\begin{align*}
    a_{11}x + a_{12}y + a_{13}z &= b_1 \\
    a_{21}x + a_{22}y + a_{23}z &= b_2 \\
    a_{31}x + a_{32}y + a_{33}z &= b_3
\end{align*}
\]

we can simply write

\[
A \, x = \bar{b}
\]

where \(\bar{b}\) is a vector containing the constants

\[
\bar{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}
\]

\(x\) is a similar vector holding the unknown quantities \((x, y\) and \(z)\), and \(A\) is a matrix, i.e. a two-dimensional grid of numbers, holding the coefficients of the unknowns.

\[
A = \begin{pmatrix}
    a_{11} & a_{12} & a_{13} \\
    a_{21} & a_{22} & a_{23} \\
    a_{31} & a_{32} & a_{33}
\end{pmatrix}
\]

Here \(A\) is a three (row) by three (column) matrix. The individual components are always referred to in the order: row then column. Thus the component \(a_{ij}\) is in the \(i\)-th row and the \(j\)-th column.

Note that in general the number of rows doesn’t have to equal the number of columns. Also, a vector is really just a special class of matrix which only has one column - you can think of a matrix as a bunch of vectors stuck together.

**Example:** How would we write the mineralogy question above in matrix format?

**Matrix Addition**

Matrix addition can only be carried out if the two matrices have the same number of rows and number of columns. In this case, the components of the summed matrix \(s_{ij}\) are simply given by

\[
s_{ij} = a_{ij} + b_{ij}
\]
**Matrix Multiplication**

Two matrices may be multiplied to produce a third. Denoting a matrix with \( r \) rows and \( c \) columns by \( r \times c \), then two matrices \( r_1 \times c_1 \) and \( r_2 \times c_2 \) can only be multiplied together if \( c_1 = r_2 \). The resulting matrix will be \( r_1 \times c_2 \).

Matrices can only be multiplied if the number of columns in the first matrix equals the number of rows in the second column. The resulting matrix has dimensions of (rows in first matrix, columns in second matrix).

Thus \( C = AB \) is a matrix with the same number of rows as \( A \) and the same number of columns as \( B \). We can specify its components by using the following formula:

\[
C_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj}
\]

where in this case \( A \) has \( n \) columns and \( B \) has \( n \) rows.

What does this mean? Here’s an example:

\[
A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad B = \begin{pmatrix} e & f \\ g & h \end{pmatrix} \quad AB = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix}
\]

What does \( BA \) equal? Does it equal \( AB \)?

**Answer:**

\[
BA = \begin{pmatrix} ea + fc & eb + fd \\ ga + hc & gb + hd \end{pmatrix}
\]

So \( BA \neq AB \).

**Example** Carry out these two multiplications:

\[
(0 \ 1 \ -1) \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}
\]

and

\[
\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} (0 \ 1 \ -1)
\]
We can also use matrix algebra to multiply (dot) two vectors (single-column matrices) together. To do this, though, we have to turn the first vector "on its side" so that the normal rules of matrix algebra are obeyed. We write the multiplication as
\[ a^T b = b^T a = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \]
Here the symbol "T" stands for Transpose, an operation which exchanges rows with columns in a vector or matrix. For example

If \( \mathbf{x} = \begin{pmatrix} x_1 & x_4 \\ x_2 & x_5 \\ x_3 & x_6 \end{pmatrix} \), then \( \mathbf{x}^T = \begin{pmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \end{pmatrix} \)

We transpose \( a \) before multiplying it with \( b \) so that the number of columns in \( a \) equals the number of rows in \( b \), as before. As before, the result is a matrix with dimensions equal to the number of rows in the first matrix (1) and the number of columns in the second (1), giving us a one-by-one matrix, i.e. a scalar. Note that in this case the order of multiplication doesn’t matter, but in general for matrices \( A \) \( B \) does not equal \( B A \).

**Special Matrices**

1. **Identity Matrix**

    Usually denoted \( I \), this is the matrix equivalent of the number 1, such that
    \[ A I = I A = A \]

    It is square (rows = columns), has 1’s along the diagonals, and zeroes everywhere else. We can specify the value of each component using Kronecker’s Delta:
    \[ \delta_{ij} = 1 \text{ if } i = j, \quad \delta_{ij} = 0 \text{ otherwise} \]

    In two dimensions, the identity matrix is given by
    \[ I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \]

2. **Diagonal Matrix**
A diagonal matrix only has non-zero values on the diagonal, but these values can be any real number (and not just 1). The identity matrix is a special kind of diagonal matrix. The components of a general diagonal matrix can be specified as $c_i \delta_{ij}$, where $c_i$ is a real number.

3. Symmetric Matrices

These are square, and distinguished by the fact that they are equal to their own transpose i.e. $A^T = A$ and $a_{ij} = a_{ji}$.

4. Upper Triangular

Any matrix for which $a_{ij} = 0$ for $i > j$. A lower triangular matrix obeys $a_{ij}=0$ for $i < j$. Triangular matrices become important when solving sets of equations.

Matrix multiplication and addition is

1. Associative: $(AB)C = A(BC)$; $(A + B) + C = A + (B + C)$

2. Distributive: $A(B + C) = AB + AC$

3. Not Commutative: $AB$ does not (usually) equal $BA$, although $A + B = B + A$ (always).

[One way of thinking about this is to recognize that matrices often represent some kind of rotation, and rotations are not generally commutative.]

There are also a couple of rules governing transpose matrices:

4. $(AB)^T = B^T A^T$

5. $(A + B)^T = A^T + B^T$

Reminders

1. No problem set this week.

2. J.P. will be holding special discussion sections (going through the midterm and practising fundamentals) this Thursday and next Wednesday

3. There are practice example questions (and answers) posted on the class web page.