Matrix Algebra

In the present section we treat the topic of matrices and matrix operations. Systems of linear equations result from many situations where we seek to describe the behaviour of a physical system. The analysis of a system of trusses for example, where we seek to determine the strains in the various strut components when the system is subject to given loads, will typically result in a system of linear equations for the strains. Similarly the analysis of currents flowing in a complex electrical circuit give rise to a system of linear equations. Matrices give a convenient means of representing such systems of equations.

We begin by considering the following simple linear system in which we wish to determine values for $\alpha$ and $\beta$ that satisfy the equations:

\[
\begin{align*}
3\alpha + 2\beta &= 7 \\
\alpha + 2\beta &= 5
\end{align*}
(1)
\]

There are two issues to concern us here. The first is does a solution to equations 1 exist? The second, if a solution exists, is it unique? The answer to both of these questions, assuming that we have formulated a valid mathematical description of a physical system, must presumably be yes. (We would also expect of course that this unique solution gives a meaningful solution to the problem we are attempting to solve.)

One solution of the above equations is $\alpha = 1, \beta = 2$ as is readily verified by substitution. Whether or not the solution is unique is not so easily resolved and will be the subject of future consideration.
In practice, problems involving systems of linear equations will consist of perhaps hundreds and even thousands of equations and variables and in order to obtain a solution we will need make use of a computer. The computer will generally store and solve the problem in terms of the numerical values of the system alone, that is the labels attached to the variables will not be stored, and so we need represent the system in such a way as to separate numbers and labels. We do this as shown in equation 2 below where we separate the components of equations 1.

\[
\begin{bmatrix}
3 & 2 \\
1 & 2 \\
\end{bmatrix}
\begin{bmatrix}
\alpha \\
\beta \\
\end{bmatrix}
= 
\begin{bmatrix}
7 \\
5 \\
\end{bmatrix} 
\] (2)

The rectangular arrays of numbers and/or symbols enclosed in square brackets in equation 2 are referred to as matrices. We will later interpret the equation in terms of a matrix algebra, that is algebraic operations performed on matrices.

The matrix with two rows and two columns on the left hand side of equation 2 is the coefficient matrix of the system, the matrix with two rows and one column on the right hand side is the right hand side matrix and the matrix with two rows and one column on the left hand side is the matrix of unknowns. In a more concise notation we usually write

\[ A \mathbf{x} = \mathbf{b} \]

where we repeat that the essential information contained in the above system is that in the matrices \( A \) and \( \mathbf{b} \). We define more carefully later how we are to interpret the ‘matrix multiplication’ \( A \mathbf{x} \) on the left hand side of equation 2. Note that the rows of the matrices \( A \) and \( \mathbf{b} \) correspond to the equations of the original system and the columns of \( A \) correspond to the variables of the system.
We generally denote the number of rows in the matrix by \( m \) and the number of columns by \( n \). We then say that we have an \( m \times n \) matrix or a \( m \) by \( n \) matrix (remember the number of rows is always given first). Thus the matrix
\[
\begin{bmatrix}
3 & 2 & 1 \\
1 & 2 & 3 \\
4 & 3 & 2 \\
3 & 4 & 5
\end{bmatrix}
\]
is a 4 by 3 matrix or a \( 4 \times 3 \) matrix.

A matrix with the same number of rows as columns is said to be a square matrix of order \( n \) (where \( n \) is the number of rows or columns) or is said simply to be square.

It is usual to refer to a matrix with a single row as a row vector and to a matrix with a single column as a column vector. Thus
\[
\begin{bmatrix}
1 \\
2 \\
3 \\
4
\end{bmatrix}
\]
are a row vector and a column vector respectively.

Also, what we defined above as the right hand side matrix and the matrix of unknowns are usually referred to as the right hand side vector and the vector of unknowns (in matrix algebra vectors are usually assumed to be column vectors unless stated otherwise).

We next consider some operations involving matrices and begin by representing the general \( m \times n \) matrix \( A \) in the form
\[
A = \begin{bmatrix}
a_{11} & a_{12} & \ldots & a_{1n} \\
a_{21} & a_{22} & \ldots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1} & a_{m2} & \ldots & a_{mn}
\end{bmatrix}
\]
where \( a_{ij} \) is the entry (or coefficient or element) in row \( i \), column \( j \) (again note we take first row and then column). Similarly the matrix \( B \) will have entries \( b_{ij} \).
We define the following matrix operations:

• Two matrices $A$ and $B$ are said to be equal if and only if they have the same number of rows and columns, and their respective entries are equal i.e $a_{ij} = b_{ij}$.

• Two matrices $A$ and $B$ may be added if they have the same number of rows and columns. The sum of the matrices is then obtained by adding the respective entries of $A$ and $B$, that is the entry in row $i$ column $j$ of $A + B$ is $a_{ij} + b_{ij}$.

• Multiplying a matrix $A$ by a scalar $\lambda$ produces a matrix with the same number of rows and columns as $A$ but with the entries of $A$ multiplied by $\lambda$, that is the entry in row $i$ column $j$ of $\lambda A$ is $\lambda a_{ij}$.

• The transpose of a matrix $A$ is denoted $A^T$ and is obtained by interchanging the rows and columns of $A$. That is, the rows of $A$ become the columns of $A^T$.

Example
The following rules governing matrix operations generally follow from the respective rules for real numbers.

- Matrix addition is commutative
  \[ A + B = B + A. \]
- Matrix addition is associative
  \[(A + B) + C = A + (B + C).\]
- Matrix multiplication by a scalar is distributive over matrix addition:
  \[ \lambda(A + B) = \lambda A + \lambda B, \]
- The transpose of the transpose a matrix \( A \) is \( A \), that is
  \[ (A^T)^T = A. \]
- The transpose of the sum of two matrices \( A \) and \( B \) is the sum of the transposed matrices, that is
  \[ (A + B)^T = A^T + B^T. \]
A number of definitions of especial use in relation to square matrices are given below.

- The diagonal of the matrix from the top left hand corner to the bottom right hand corner is the main diagonal or (principal diagonal or leading diagonal).
- The sum of the entries on the main diagonal is the trace of the matrix.
- A diagonal matrix is a matrix whose entries off the main diagonal are all zero.
- An upper triangular matrix is a matrix whose entries below the main diagonal are all zero.
- A lower triangular matrix is a matrix whose entries above the main diagonal are all zero.
- A diagonal matrix whose entries on the main diagonal are all one is a unit matrix (or unit matrix of order \( n \)) or an identity matrix (identity matrix of order \( n \)). Such a matrix is denoted \( I \) or \( I_n \) if the order needs be denoted explicitly.
- A (square) matrix \( A \) for which \( A = A^T \) is said to be a symmetric matrix.
- A (square) matrix \( A \) for which \( A = -A^T \) is said to be anti-symmetric or skew symmetric.

We repeat that the above definitions are applicable to square matrices only.

Example
It remains to discuss the most useful operation involving matrices: *matrix multiplication.*

We came across matrix multiplication when we expressed a system of equations in the matrix form $A\mathbf{x} = \mathbf{b}$. Below the original system of equations is given on the left, and the system in matrix notation on the right.

\[
\begin{align*}
3\alpha + 2\beta &= 7 \\
\alpha + 2\beta &= 5
\end{align*}
\]

\[
\begin{bmatrix}
3 & 2 \\
1 & 2
\end{bmatrix}
\begin{bmatrix}
\alpha \\
\beta
\end{bmatrix}
= 
\begin{bmatrix}
7 \\
5
\end{bmatrix}
\]

We clearly must define matrix multiplication in such a way that the system on the right is equivalent system of equations on the left.

We begin by defining the product of a row matrix with a column matrix. Consider the definition of the product

\[
\begin{bmatrix}
a & b & c & d & e
\end{bmatrix}
\begin{bmatrix}
f \\
g \\
h \\
i \\
j
\end{bmatrix}
= a \times f + b \times g + c \times h + d \times i + e \times j
\]

That is we pair the entries in the vectors, multiply the pairs and add. We will refer to this operation as ‘multiplication of a row into a column.’ It is evident that the operation only makes sense if the number of entries in the row is the same as the number of entries in the column. Taking a numerical example we have

\[
\begin{bmatrix}
1 & 2 & 3 & 4
\end{bmatrix}
\begin{bmatrix}
5 \\
6 \\
7 \\
8
\end{bmatrix}
= 1 \times 5 + 2 \times 6 + 3 \times 7 + 4 \times 8 = 5 + 12 + 21 + 32 = 70.
\]

For the general case, where we multiply vectors with $n$ entries, we write

\[
\begin{bmatrix}
a_1 & a_2 & a_3 & \ldots & a_n
\end{bmatrix}
\begin{bmatrix}
b_1 \\
b_2 \\
b_3 \\
\vdots \\
b_n
\end{bmatrix}
\]
\[ a_1 \times b_1 + a_2 \times b_2 + a_3 \times b_3 + \ldots + a_n \times b_n. \]

We briefly note the fact here (this will be expanded on later) that the above can be viewed as an extension of the vector scalar product applied to vectors with more than three entries. In the present context this multiplication of a row vector into a column vector is usually referred to as an inner product.

We next extend the above process to the multiplication of a matrix into a column vector. An example is

\[
\begin{bmatrix}
1 & 2 & 3 & 4 \\
5 & 6 & 7 & 8 \\
9 & 10 & 11 & 12
\end{bmatrix}
\begin{bmatrix}
5 \\
6 \\
7 \\
8
\end{bmatrix}
\]

In this case we multiply rows of the first into the column of the second. Since there are three rows there are three possible products of rows into columns and each product gives a row entry in the resultant column vector.

\[
\begin{bmatrix}
1 & 2 & 3 & 4 \\
5 & 6 & 7 & 8 \\
9 & 10 & 11 & 12
\end{bmatrix}
\begin{bmatrix}
5 \\
6 \\
7 \\
8
\end{bmatrix}
= \begin{bmatrix}
1 \times 5 + 2 \times 6 + 3 \times 7 + 4 \times 8 \\
5 \times 5 + 6 \times 6 + 7 \times 7 + 8 \times 8 \\
9 \times 5 + 10 \times 6 + 11 \times 7 + 12 \times 8
\end{bmatrix}
= \begin{bmatrix}
70 \\
174 \\
278
\end{bmatrix}
\]

Note here that we are multiplying a matrix with three rows and four columns into a matrix with four rows and one column. The result is a matrix with three rows and one column. The rule is

\[(3 \times 4)(4 \times 1) \rightarrow (3 \times 1)\]

where the number or columns of the first matrix must be the same as the number of columns of the second - here both are equal to three. The general situation when multiplying a general matrix into a column vector is

\[(m \times p)(p \times 1) \rightarrow (m \times 1).\]
It only remains to generalise matrix multiplication to the case

$$(m \times p)(p \times n) \rightarrow (m \times n).$$

We are still multiplying rows of the first matrix into columns of the second matrix. We can do this as the number of columns of the first matrix is the same as the number of rows of the second (both $p$). The entry in row $i$ column $j$ of the resulting matrix is obtained by multiplying row $i$ of the first matrix into column $j$ of the second. Thus if we multiply $AB = C$ we have

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \ldots + a_{ip}b_{pj} = \sum_{k=1}^{p}a_{ik}b_{kj}$$

**Example** Consider the multiplication of matrices $A$ and $B$ where

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix}$$

We can form the product $AB$:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 \times 1 + 2 \times 2 + 3 \times 3 & 1 \times 2 + 2 \times 3 + 3 \times 4 \\ 4 \times 1 + 5 \times 2 + 4 \times 3 & 4 \times 2 + 5 \times 3 + 4 \times 4 \end{bmatrix}$$

$$= \begin{bmatrix} 1 + 4 + 9 & 2 + 6 + 12 \\ 4 + 10 + 18 & 8 + 15 + 24 \end{bmatrix} = \begin{bmatrix} 14 & 20 \\ 32 & 47 \end{bmatrix}$$

and the product $BA$:

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 1 \times 1 + 2 \times 4 & 1 \times 2 + 2 \times 5 & 1 \times 3 + 2 \times 6 \\ 2 \times 1 + 3 \times 4 & 2 \times 2 + 3 \times 5 & 2 \times 3 + 3 \times 6 \\ 3 \times 1 + 4 \times 4 & 3 \times 2 + 4 \times 5 & 3 \times 3 + 4 \times 6 \end{bmatrix}$$

$$= \begin{bmatrix} 1 + 8 & 2 + 10 & 3 + 12 \\ 2 + 12 & 4 + 15 & 6 + 18 \\ 3 + 16 & 6 + 20 & 9 + 24 \end{bmatrix} = \begin{bmatrix} 14 & 19 & 24 \\ 19 & 26 & 33 \end{bmatrix}$$

It is clear that $AB \neq BA$. That is, matrix multiplication is not commutative.
We mentioned previously that the multiplication of a row vector into a
column vector is often termed an inner product. To be precise the inner
product is usually taken to be between two column vectors. The inner
product of column vectors $\mathbf{x}$ and $\mathbf{y}$ is in fact $\mathbf{x}^T \mathbf{y}$ and is usually taken to be
a real number rather than a one by one matrix (in most situations a one by
one matrix is simply replaced by a real number).

We now list some properties of matrix multiplication. The properties are
generally easily proved though the proofs are sometimes tedious. It is implicitly assumed below that the operations of matrix addition and multiplication are between compatible matrices, that is the operations are defined.

- Matrix multiplication is not commutative. That is in general
  \[ AB \neq BA. \]

- Matrix multiplication is associative. That is, if the product
  \((AB)C\) is defined, then
  \[ (AB)C = A(BC). \]

- For multiplication by a scalar $\lambda$
  \[ (\lambda A)B = \lambda(AB). \]

- Matrix multiplication is distributive over matrix addition of matrices:
  \[ A(B + C) = AB + AC; \]
  \[ (A + B)C = AC + BC. \]

- If the matrix $A$ is $m \times n$ and $I_n$ denotes the square matrix of
  order $n$ then
  \[ AI_n = I_m A = A \]

- For the transpose of the product of matrices we have
  \[ (AB)^T = B^T A^T \]
  that is the transpose of a product is the product of the transposed
  matrices but in the reverse order.