The idea of a matrix

In mathematics and its applications, we frequently consider functions which take a family of real numbers $x_1, ..., x_n$ to another family of real numbers $y_1, ..., y_m$. The simplest functions of this sort are those in which each y_i is expressed as a linear combination of the x_j's; i.e., as an expression of the form $a_1 x_1 + ... + a_n x_n$. Since we need a different such expression for each of the terms $y_1, ..., y_m$ of our output-string, we use double subscripts, writing the formula for the ith term as

$$y_i \;=\; a_{i1} x_1 + ... + a_{in} x_n .$$

Our function is then determined by the mn coefficients a_{ij}; hence we may use the array of these coefficients:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

as an abbreviation for the function. Writing the input- and output-families as column vectors, this function may be described by the formula:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \;=\; \begin{bmatrix} a_{11} x_1 + ... + a_{1n} x_n \\ \vdots \\ a_{m1} x_1 + ... + a_{mn} x_n \end{bmatrix}.$$

An array of numbers such as we have formed from the coefficients a_{ij} is called an ***m×n matrix***. You will see many apparently different uses of matrices in your mathematics courses, but essentially all of them can be explained as instances of the functional interpretation given above.

The set of all n-tuples of real numbers is commonly denoted $\mathbb{R}^n$, so the functions discussed above are maps $\mathbb{R}^n \to \mathbb{R}^m$. A map of the form we described in the first paragraph is called a ''linear transformation''. In summary

$m \times n$ matrices of real numbers represent linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$.

(The same is true with integers, rational numbers, or complex numbers in place of real numbers, in which case we have $\mathbb{Z}^n$, $\mathbb{Q}^n$, $\mathbb{C}^n$ etc. in place of $\mathbb{R}^n$ etc..)

The *matrix operations* all have natural interpretations in terms of the above interpretation of matrices. Thus, if A and B are $m \times n$ matrices representing two linear transformations $\mathbb{R}^n \to \mathbb{R}^m$, then $A+B$ is the $m \times n$ matrix representing the transformation that takes each vector $\mathbf{x} \in \mathbb{R}^n$ to the *sum* of the vectors $A\mathbf{x} \in \mathbf{R}^m$ and $B\mathbf{x} \in \mathbf{R}^m$:

$$(A+B)\mathbf{x} \;=\; A\mathbf{x} + B\mathbf{x} .$$

Multiplication of matrices, on the other hand, corresponds to *composition* of linear transformations. That is, given three positive integers $m, n, p,$ and linear transformation $\mathbb{R}^p \to \mathbb{R}^n$ and $\mathbb{R}^n \to \mathbb{R}^m$, represented respectively by an $n \times p$ matrix B and an $m \times n$ matrix A, the composite map $\mathbb{R}^p \to \mathbb{R}^m$, i.e., the map taking each vector $\mathbf{x} \in \mathbf{R}^p$ to the vector $A(B\mathbf{x}) \in \mathbb{R}^m$, will also be linear, and the $m \times p$ matrix which represents it is denoted AB, so that we have the law

$$(AB)\mathbf{x} \;=\; A(B\mathbf{x}).$$

You can check that the matrix AB which satisfies the above equation is indeed the one given by the rule in Fraleigh for multiplying matrices; that is, its entry in the ith row and jth column is obtained from the ith row of A and the jth column of B by formula (9) on p.27.