The joy of matrix maths

Matrices

A matrix is a rectangle of numbers. For example:

$$X = \begin{pmatrix} 1 & 3 & 2 \\ 4 & 2 & 7 \end{pmatrix}$$

X is a matrix with 2 rows and 3 columns, a 2 x 3 matrix. In general a matrix with a rows and b columns is an a x b matrix.

The individual numbers in the matrix are referred to as elements. If we want to refer to a particular element, we can do so using its row and column number. For example, we could refer to the element in the second row and the third column of A as a_{23} - in this case a_{23} is 7.

A vector is a matrix with only one row, or only one column. A row vector has only one row - e.g:

$$R = \begin{pmatrix} 1 & 3 & 2 & 4 & 2 & 7 \end{pmatrix}$$

and a column vector has only one column:

$$C = \begin{pmatrix} 1 \\ 3 \\ 2 \\ 4 \\ 2 \\ 7 \end{pmatrix}$$

A scalar is a matrix with one row and one column - a 1 x 1 matrix - a single number.

Adding matrices

Adding matrices is easy. In order to add two matrices, one of two things must be true. Either the two matrices must have the same number of rows, and the same number of columns, or one of the matrices must be a scalar. Thus, if you have an a x b matrix, and you want to add this to another matrix, that matrix must either be an a x b matrix, or a scalar.

To add a scalar to a matrix, you merely add the scalar value to all the elements in the matrix:

$$\begin{pmatrix} 1 & 3 & 2 \\ 4 & 2 & 7 \end{pmatrix} + 3 = \begin{pmatrix} 4 & 6 & 5 \\ 7 & 5 & 10 \end{pmatrix}$$

To add an a x b matrix to another a x b matrix matrix, you take each of the elements in the first matrix, and add to it the element from the second matrix that is in the same row and the same column. So, say we are adding two matrices, A and B, to get the result matrix C. The element in the first row and the first column of C (c_{11}) is a_{11} + b_{11}; the element in the first row and the second column of C (c_{12}) is a_{12} + b_{12}, and the element in row i and column j of C is a_{ij} + b_{ij}. Thus:

$$\begin{pmatrix} 1 & 3 & 2 \\ 4 & 2 & 7 \end{pmatrix} + \begin{pmatrix} 2 & 1 & 1 \\ 6 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1+2 & 3+1 & 2+1 \\ 4+6 & 2+3 & 7+2 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 3 \\ 10 & 5 & 9 \end{pmatrix}$$

Multiplying matrices

It is very easy to multiply a matrix by a scalar. Like addition, you merely multiply all the elements in the matrix by the scalar:

$$\begin{pmatrix} 1 & 3 & 2 \\ 4 & 2 & 7 \end{pmatrix} * 3 = \begin{pmatrix} 3 & 9 & 6 \\ 12 & 6 & 21 \end{pmatrix}$$

Multiplying matrices is just a little less straightforward. In fact, matrix multiplication - when neither matrix is a scalar - has some rather odd properties:

> If A is a matrix and B is a matrix, A * B may well not be the same as B * A
>
> In order for it to be possible to determine the result of A * B, A must have the same number of columns that B has rows

To repeat: if you have a matrix A that has a rows and b columns (an a x b matrix), and another matrix B that is c x d, then getting a result of A * B is only possible if $b = c$. In fact, the result of A * B is a new matrix of size a x d.

So, to start with the simplest case - multiplying two vectors. Let us say we have a row vector:

$$R = \begin{pmatrix} 1 & 3 & 2 & 4 & 2 & 7 \end{pmatrix}$$

If we are going to multiply R by another vector, C (i.e. R * C), then C *must* either be a scalar (see above) or a column vector, with the same number of elements as R. So, let us say that we have such a column vector:

$$C = \begin{pmatrix} 8 \\ 1 \\ 6 \\ 3 \\ 9 \\ 5 \end{pmatrix}$$

The result of R * C is a new matrix, S, and in fact S is a scalar (the result of a 1x6 * 6x1 multiplication is 1x1). The reason that S is a scalar is because the multiplication works like this: each element in R is multiplied by the equivalent element in C. So, you take the first element in R and multiply it by the first element in C, then you take the second element in R and multiply it by the second element in C, and so on. If R is a 1 x n vector (so C must be a n by 1 vector), then this gives us n multiplication results. To get the final value of the vector multiplication, you sum the results of the element by element multiplications. So, for example:

$$\begin{pmatrix} 1 & 3 & 2 & 4 & 2 & 7 \end{pmatrix} * \begin{pmatrix} 8 \\ 1 \\ 6 \\ 3 \\ 9 \\ 5 \end{pmatrix} = \left([1 * 8] + [3 * 1] + [2 * 6] + [4 * 3] + [2 * 9] + [7 * 5]\right) = 88$$

Matrix multiplication is just an extension of vector multiplication. Let's say we want to multiply two matrices - A * B, to get a result C:

$$A = \begin{pmatrix} 1 & 3 & 2 \\ 4 & 2 & 7 \end{pmatrix} B = \begin{pmatrix} 6 & 3 \\ 1 & 2 \\ 8 & 7 \end{pmatrix}$$

We can do this by breaking the matrices up into vectors. We break A up into *row* vectors, and B up into *column* vectors. We then take each of the row vectors of A in turn, and multiply them by all the column vectors in B. So, to get the element in the the first row and first column of C, we vector multiply the first row in A, by the first column in B:

$$c_{11} = \begin{pmatrix} 1 & 3 & 2 \end{pmatrix} * \begin{pmatrix} 6 \\ 1 \\ 8 \end{pmatrix} = \Big([1 * 6] + [3 * 1] + [2 * 8]\Big) = 25$$

Then we multiply the first row in A by the second column in B, to get a value for C that is in the first row, and second column:

$$c_{12} = \begin{pmatrix} 1 & 3 & 2 \end{pmatrix} * \begin{pmatrix} 3 \\ 2 \\ 7 \end{pmatrix} = \Big([1 * 3] + [3 * 2] + [2 * 7]\Big) = 23$$

When we have finished multiplying the first row of A by the columns of B, we take the second row of A, and multiply by the columns of B again. This time the results go into the *second row* of C:

$$c_{21} = \begin{pmatrix} 4 & 2 & 7 \end{pmatrix} * \begin{pmatrix} 6 \\ 1 \\ 8 \end{pmatrix} = \Big([4 * 6] + [2 * 1] + [7 * 8]\Big) = 82$$

$$c_{22} = \begin{pmatrix} 4 & 2 & 7 \end{pmatrix} * \begin{pmatrix} 3 \\ 2 \\ 7 \end{pmatrix} = \Big([4 * 3] + [2 * 2] + [7 * 7]\Big) = 65$$

So:

$$\begin{pmatrix} 1 & 3 & 2 \\ 4 & 2 & 7 \end{pmatrix} * \begin{pmatrix} 6 & 3 \\ 1 & 2 \\ 8 & 7 \end{pmatrix} = \begin{pmatrix} 25 & 23 \\ 82 & 65 \end{pmatrix}$$