Matrices and Linear Algebra

Let's say that there is a black box in which if you enter three numbers (x_1, x_2, x_3), two numbers (y_1, y_2) pop out. (If you know what a vector is, you can regard the three numbers as a three-dimensional vector and the two numbers as a two-dimensional vector.) For example, let's say that you entered (1, 2, 0) for (x_1, x_2, x_3), and you got (1, 4) for (y_1, y_2).

Now, let's say that you found out, by playing with this black box, that it obeys the following properties:

Firstly, if you enter the numbers that are n times as big as the original numbers, the black box pops out the numbers that are n times as big as the original output.

For example, let's use the example above as a base to demonstrate this property. If you enter (2, 4, 0) which are two times as big as (1, 2, 0), you get (2, 8) which are two times as big as (1, 4). One more example: If you enter (3, 6, 0) which are three times as big as (1, 2, 0), you get (3, 12) which are three times as big as (1, 4).

Then, which numbers would this black box pop out, if you entered (0, 0, 0) for (x_1, x_2, x_3)?

It will pop out (0, 0) as (0, 0, 0) is (1, 2, 0) multiplied by 0 and (0, 0) is (1, 4) multiplied by 0.

Secondly, if you enter $(a_1 + b_1, a_2 + b_2, a_3 + b_3)$ for (x_1, x_2, x_3), the black box pops out numbers which are the sum of the numbers that would have been popped out if (a_1, a_2, a_3) were entered, and the numbers that would have been popped out if (b_1, b_2, b_3) were entered. For example, if you entered (1, 2, 0) and, (1, 4) were popped out, and if you entered (3, 4, -1) and (4, 6) were popped out, then (5, 10) would have been popped out if (4=1+3. 6=2+4, -1=0-1) were entered.

A black box, or rather a function which satisfies the above conditions called "linearity" is called "a linear function" or "a linear operator." (Notice that a linear operator is a function from a vector to a vector. In our case, from a three-dimensional vector to a two-dimensional vector.)

If we write this function in terms of the mathematical language, we can write it as:

$$f(x_1, x_2, x_3) \ = (y_1, y_2).$$

For examples:

$$f\,(1,\ 2,\ 0) = (1,\ 4)$$

$$f\,(3,\ 4,\ -1) = (4,\ 6)$$

Let me write the conditions the black box obeys in a mathematical language.

The first property can be written as follows:

$$f\left(nx_1, nx_2, nx_3\right) = \left(ny_1, ny_2\right)$$

The second property can be written as follows:

$$f\left(a_1 + b_1, a_2 + b_2, a_3 + b_3\right) = f\left(a_1, a_2, a_3\right) + f(b_1, b_2, b_3)$$

Then, let me ask you the following question.

If we knew that this black box was a linear function or a linear operator beforehand, how many operations would have been needed to completely predict the operation of this black box?

The answer is three.

The two properties of the linear operator are very important to understand this answer. By using these properties, we can write as follows:

$$\begin{aligned} f(x_1, x_2, x_3) = f(x_1 + 0, x_2 + 0, 0 + x_3) = f(x_1, x_2, 0) + f(0, 0, x_3) \\ = f(x_1 + 0, 0 + x_2, 0 + 0) + f(0, 0, x_3) = f(x_1, 0, 0) + f(0, x_2, 0) + f(0, 0, x_3) \end{aligned}$$

where we have used the second property of the linear operator.

Now, by using the first property, we can write the following:

$$\begin{aligned} f(x_1, x_2, x_3) &= f(x_1, 0, 0) + f(0, x_2, 0) + f(0, 0, x_3) \\ &= x_1 f(1, 0, 0) + x_2 f(0, 1, 0) + x_3 f(0, 0, 1) \end{aligned}$$

In other words, the value of $f(x_1, x_2, x_3)$ is completely determined once we know the values of $f(1, 0, 0)$, $f(0, 1, 0)$, and $f(0, 0, 1)$.

Now, let's say that the value for $f(1, 0, 0)$is (f_{11}, f_{21}), and the value for $f(0, 1, 0)$ is (f_{12}, f_{22}), and the value for $f(0, 0, 1)$ is (f_{13}, f_{23}).

Then, we can write as follows:

$$\begin{aligned} f\left(x_1, x_2, x_3\right) = \ x_1 \ \left(f_{11},\ f_{21}\right) + \ x_2 \ \left(f_{12},\ f_{22}\right) + x_3 \ \left(f_{13},\ f_{23}\right) \\ = \left(x_1 f_{11} + x_2 f_{12} + x_3 f_{13},\ x_1 f_{21} + x_2 f_{22} + x_3 f_{23}\right) \end{aligned}$$

If we take advantage of the notation called "matrix," we can write the above equation as follows:

$$\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

Of course, y_1 and y_2 are given by "$x_1 f_{11} + x_2 f_{12} + x_3 f_{13}$," and "$x_1 f_{21} + x_2 f_{22} + x_3 f_{23}$."

In other words, the above equation defines the matrix multiplication. When I first learned the rule for the matrix multiplication, I didn't understand why it was defined this way, as I didn't know that matrix is just an easy way of writing a linear operator. Now, I hope that you are more motivated.

Again, if we call "$\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \end{bmatrix}$" "$F$," "$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$" "$X$," and "$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$" "$Y$," we can write the above equation as:

$$FX = Y \tag{1}$$

Or, if we use another notation, we can write it as:

$$\sum_{b=1}^{3} f_{ab} x_b = y_a$$

Matrices which have two rows and three columns such as F are called "2×3" matrices. Matrices which have three rows and one column such as X are called "3×1" matrices. Similarly, Y is a "2×1" matrix.

The matrix multiplication is possible in more general cases. In other words, it isn't necessary that X and Y have single columns in the above equation $FX = Y$. To obtain a definition for more general matrix multiplications, let's consider the case that two different linear operators successively act on X.

For example, let g be a linear operator, which pops out four numbers if two numbers are entered.
In other words,

$$g(y_1, y_2) = \ (z_1, z_2, z_3, z_4)$$

Now, as I said, let's consider the following composition of two linear operators g and f.

$$g\left(f\left(x_1, x_2, x_3\right)\right) = \ g(y_1, y_2) = \ (z_1, z_2, z_3, z_4)$$

If we use the notation of matrix, we can write as follows:

$$\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ g_{31} & g_{32} \\ g_{41} & g_{41} \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ g_{31} & g_{32} \\ g_{41} & g_{41} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix}$$

In other words,

$$GFX = GY = Z \tag{2}$$

If we say H is GF, a linear operator which pops out four numbers when three numbers were entered, we can obtain h by the following way:

$$\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ g_{31} & g_{32} \\ g_{41} & g_{41} \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \\ g_{31} & g_{32} \\ g_{41} & g_{41} \end{bmatrix} \begin{bmatrix} f_{11}x_1 + f_{12}x_2 + f_{13}x_3 \\ f_{21}x_1 + f_{22}\ x_2 + f_{23}x_3 \end{bmatrix}$$

$$= \begin{bmatrix} g_{11}\,(f_{11}x_1 + f_{12}x_2 + f_{13}x_3) + g_{12}(f_{21}x_1 + f_{22}\ x_2 + f_{23}x_3) \\ g_{21}\,(f_{11}x_1 + f_{12}x_2 + f_{13}x_3) + g_{22}(f_{21}x_1 + f_{22}\ x_2 + f_{23}x_3) \\ g_{31}\,(f_{11}x_1 + f_{12}x_2 + f_{13}x_3) + g_{32}(f_{21}x_1 + f_{22}\ x_2 + f_{23}x_3) \\ g_{41}\,(f_{11}x_1 + f_{12}x_2 + f_{13}x_3) + g_{42}(f_{21}x_1 + f_{22}\ x_2 + f_{23}x_3) \end{bmatrix}$$

$$= \begin{bmatrix} g_{11}f_{11} + g_{12}f_{21} & g_{11}f_{12} + g_{12}f_{22} & g_{11}f_{13} + g_{12}f_{23} \\ g_{21}f_{11} + g_{22}f_{21} & g_{21}f_{12} + g_{22}f_{22} & g_{21}f_{13} + g_{22}f_{23} \\ g_{31}f_{11} + g_{32}f_{21} & g_{31}f_{12} + g_{32}f_{22} & g_{31}f_{13} + g_{32}f_{23} \\ g_{41}f_{11} + g_{42}f_{21} & g_{41}f_{12} + g_{42}f_{22} & g_{41}f_{13} + g_{42}f_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \\ h_{41} & h_{42} & h_{43} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

In other words,

$$GFX = G(FX) = (GFX) = (GF)X = HX \tag{3}$$

If we use another notation:

$$\sum_{a=1}^{2}\sum_{b=1}^{3} g_{ca}f_{ab}x_b = z_c = \sum_{b=1}^{3} h_{cb}x_b$$

Therefore, we obtain:

$$\sum_{a=1}^{2}\sum_{b=1}^{3} g_{ca}f_{ab} = z_c = \sum_{b=1}^{3} h_{cb}$$

Therefore, if G is a 4×2 matrix and F is a 2×3 matrix, then $H = GF$ is a 4×3 matrix and its value is given by the above equation.

Generally speaking if "A" is a "$m \times p$" matrix, "B" a "$p \times n$" matrix, then "AB" is a "$m \times n$" matrix.

This condition must be satisfied for matrix A and matrix B to be multiplied together. For example if "A" is a "3×4 matrix", and B a "5×3 matrix, " AB is not defined as 4 is not equal to 5.

Another interesting fact is that AB is not equal to BA, for two matrices A and B, unlike in the case of ordinary numbers. For example, let's say:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \tag{4}$$

$$AB = \begin{bmatrix} 1 \times 5 + 2 \times 7 & 1 \times 6 + 2 \times 8 \\ 3 \times 5 + 4 \times 7 & 3 \times 6 + 4 \times 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix} \tag{5}$$

However,

$$BA = \begin{bmatrix} 5 \times 1 + 6 \times 3 & 5 \times 2 + 6 \times 4 \\ 7 \times 1 + 8 \times 3 & 7 \times 2 + 8 \times 4 \end{bmatrix} = \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix} \tag{6}$$

This fact is very important in quantum mechanics. According to physicists, there are corresponding linear operators, or, equivalently, matrices, for every observable such as position, momentum, and energy. Let's call the matrix corresponding to position "X", and the matrix corresponding to momentum "P."(The momentum is defined by the velocity multiplied by the mass.) Then, according to quantum mechanics, we obtain XP not equal to PX. Heisenberg derived his uncertainty principle from this observation. Heisenberg's uncertainty principle says that the more precisely you try to measure the position of an object, the less precisely you can measure the momentum of the object and vice versa.

Finally, the subject which deals with matrices or equivalently linear operators is called "linear algebra." Linear algebra is the basis of quantum mechanics. Moreover, linear algebra plays vital roles in virtually all the fields of natural science and engineering. Therefore, most of the college students studying natural science or engineering take linear algebra in their second semester of freshman year.

Problem 1. Calculate the following.

$$\begin{bmatrix} 1 & 0 & 2 \\ 3 & 0 & 4 \\ 1 & -1 & 0 \\ -3 & 0 & 0 \end{bmatrix} \begin{bmatrix} a & d \\ b & e \\ c & f \end{bmatrix} =?$$

Problem 2. Express a linear operator T in terms of matrix if it acts on a vector as follows:

$$T \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 2v_2 \\ v_1 \\ 3v_3 \\ v_1 + v_2 \end{bmatrix}$$

Problem 3. If we have A and B, two linear operators of same size (i.e. same numbers of column and same numbers of row), we can define their sum $C = A + B$ by following definition.

$$Cx = Ax + Bx \tag{7}$$

where x are the numbers (i.e. a vector) we enter into the linear operator. In other words, this equation gives what C spits out when x is entered. Convince yourself, in such a case, component-wise the following is true:

$$C_{ij} = A_{ij} + B_{ij} \tag{8}$$

If

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \tag{9}$$

what is C?

Problem 4. Similarly, we can multiply a matrix by a scalar. (i.e. a number) For example, if we have a matrix D and a scalar c, and if their product is $G = cD$, we have:

$$Gx = c(Dx) \tag{10}$$

Convince yourself, component-wise, we have the following:

$$G_{ij} = cD_{ij} \tag{11}$$

Using (8) and (11), calculate the following:

$$3\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} =? \tag{12}$$

Problem 5. A matrix A is called "diagonal matrix," if the only non-zero elements are in the diagonal part. (i.e. the top left to bottom right diagonal part) In other words, $A_{ij} = 0$ if $i \neq j$. Here are some examples of diagonal matrices:

$$\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \tag{13}$$

Here are some examples of matrices that are *not* diagonal.

$$\begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{bmatrix} \tag{14}$$

Convince yourself that two $n \times n$ diagonal matrices A and B satisfy $AB = BA$ by explicit calculations. This fact was used when Murray Gell-Mann proposed his quark model.[1]

[1] See our earlier article "Paulis exclusion principle, Color Charge of Quarks, Asymptotic freedom and Confinement" to remind yourself what the quark model was.