Matrix Multiplication

I. A Matrix times a column

To see why matrix multiplication is defined the way it is, let's start with a single linear equation in, say, 4 variables:

$$ax + by + cz + dw = f.$$

Note first that the left-hand side can be written as a dot product: $ax + by + cz + dw = (a, b, c, d) \cdot (x, y, z, w)$. We can write the equation as

$$(a, b, c, d) \cdot (x, y, z, w) = f.$$

It is customary in linear algebra to *write vectors as columns.* This has the downside of taking up a lot of room, but its advantages will soon become clear. Here is the form of our equation, with the vector of variables expressed as a column:

$$\underbrace{\begin{pmatrix} a & b & c & d \end{pmatrix}}_{A} \cdot \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = (f).$$

We have labeled the row (a, b, c, d) with an A, thinking of it as a *matrix* with 1 row and 4 columns; i.e., a 1×4 matrix. Thus, to multiply a 1×4 matrix by a 4×1 column, we simply take the dot product, obtaining a single number—the 1×1 matrix (f).

Now suppose we have two equations in these same variables:

$$\begin{aligned} ax + by + cz + dw &= f \\ a'x + b'y + c'z + d'w &= f'. \end{aligned}$$

The matrix of coefficients A now can be written with *two* rows:

$$A = \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \end{pmatrix}.$$

The two left-hand sides of our equations are the dot products of the two rows of A by the same column of variables, and the two separate results of these dottings are put into a matrix with one column and two rows:

$$\underbrace{\begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \end{pmatrix}}_{A} \cdot \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} (a, b, c, d) \cdot (x, y, z, w) \\ (a', b', c', d') \cdot (x, y, z, w) \end{pmatrix}$$

which we set equal to the column of constants:

$$A \cdot \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} (a, b, c, d) \cdot (x, y, z, w) \\ (a', b', c', d') \cdot (x, y, z, w) \end{pmatrix} = \begin{pmatrix} f \\ f' \end{pmatrix}.$$

Thus, we get a first rule of matrix multiplication:

To multiply a matrix by a column, dot each row of the matrix by the column, and put the results into a column matrix.

Example 1. Express these equations in matrix form:

$$\begin{aligned} 2x - 5y + 7z + u + 2w &= 8 \\ x + 7y + 3z - 4u + w &= 19 \\ -x + 8y \quad + \quad 2u - 9w &= -3. \end{aligned}$$

Solution:

$$\begin{pmatrix} 2 & -5 & 7 & 1 & 2 \\ 1 & 7 & 3 & -4 & 1 \\ -1 & 8 & 0 & 2 & -9 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ u \\ w \end{pmatrix} = \begin{pmatrix} 8 \\ 19 \\ -3 \end{pmatrix}.$$

In order for these dot products to make sense, you have to be able to dot the rows of A with the column (of variables). Each row of A has k entries, one for each column, so the column (of variables) must also have k entries. We therefore modify our rule for matrix multiplication.

To multiply an $n \times k$ matrix by a $k \times 1$ column:
Dot each row of the matrix by the column, and put these n results into an $n \times 1$ column.

Example 2. Calculate these matrix products if possible:

$$\begin{pmatrix} 2 & 5 & 7 \\ 0 & 1 & -6 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ -3 \\ 4 \end{pmatrix} \text{ and } \begin{pmatrix} 2 & 5 & 7 \\ 0 & 1 & -6 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix} \text{ and } \begin{pmatrix} 2 & -5 & 7 & 1 & 2 \\ 1 & 7 & 3 & -4 & 1 \\ -1 & 8 & 0 & 2 & -9 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \\ 2 \end{pmatrix}.$$

Solution. The first of these products is impossible, since the rows have 3 entries while the column has 4: you can't take the dot product. In the case of the second, $(2, 5, 7) \cdot (2, 1, -3) = -12$, $(0, 1, -6) \cdot (2, 1, -3) = 19$, and $(0, 1, 1) \cdot (2, 1, -3) = -2$; thus, the answer is the column $\begin{pmatrix} -12 \\ 19 \\ -2 \end{pmatrix}$. The third product is also defined, and the answer is $\begin{pmatrix} 17 \\ 26 \\ -3 \end{pmatrix}$.

For theoretical purposes it is sometimes useful to express the product of a matrix by a column vector in the general case. We start with an $n \times k$ matrix A and a $k \times 1$ column $\mathbf{C}$. We denote by a_{ij} the entry which lies in the ith row and jth column of A. (In describing matrices, we always give the *row information first, the column last.*) Thus, the ith row of A is the row vector

$$(a_{i1}, a_{i2}, a_{i3}, \ldots, a_{ik}).$$

If we dot this with the column $\mathbf{C}$ whose entries are $c_1, c_2, c_3, \ldots, c_k$ we get the number $p_i = \sum_{j=1}^{k} a_{ij} c_j$ and the product $A \cdot \mathbf{C}$ is the column made up of these p_i. Symbolically we have:

$$A \cdot \mathbf{C} = (a_{ij}) \cdot (c_i) = \left(\sum_{j=1}^{k} a_{ij} c_j \right).$$

II. Two ways of looking at a matrix times a column

There are two ways of looking at the product $A \cdot \mathbf{C}$ of a matrix by a column. The first way is the one we just saw:

$A{\cdot}\mathbf{C}$ = a column whose entries are the dot products of the rows of A with $\mathbf{C}$.

The other way is to see what multiplying A by $\mathbf{C} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}$ does to the *columns* of A. So, we write $A = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ where each $\mathbf{v}_j$ is an $n \times 1$ column. We have:

$$
\begin{aligned}
A \cdot B = (a_{i\ j}) \cdot (c_i) = \begin{pmatrix} \sum_{j=1}^{k} a_{1j}c_j \\ \sum_{j=1}^{k} a_{2j}c_j \\ \vdots \\ \sum_{j=1}^{k} a_{nj}c_j \end{pmatrix} &= \begin{pmatrix} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1k}c_k \\ a_{21}c_1 + a_{22}c_2 + \cdots + a_{2k}c_k \\ \vdots \\ a_{n1}c_1 + a_{n2}c_2 + \cdots + a_{nk}c_k \end{pmatrix} \\
&= \left(c_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + c_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} + \cdots + c_k \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{pmatrix} \right) \\
&= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k.
\end{aligned}
$$

This gives the other description of the matrix product:

$A{\cdot}\mathbf{C}$ = a linear combination of the columns of A with coefficients the entries of C.

Example 3.(See Example 1 above).

$$
\begin{aligned}
&\begin{pmatrix} 2 & -5 & 7 & 1 & 2 \\ 1 & 7 & 3 & -4 & 1 \\ -1 & 8 & 0 & 2 & -9 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ u \\ w \end{pmatrix} \\
&= x \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} + y \begin{pmatrix} -5 \\ 7 \\ 8 \end{pmatrix} + z \begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix} + u \begin{pmatrix} 1 \\ -4 \\ 2 \end{pmatrix} + w \begin{pmatrix} 2 \\ 1 \\ -9 \end{pmatrix} \\
&= \begin{pmatrix} 2x - 5y + 7z + u + 2w \\ x + 7y + 3z - 4u + w \\ -x + 8y \quad + \quad 2u - 9w \end{pmatrix}.
\end{aligned}
$$

Example 4. (See Example 2 above).

$$
\begin{aligned}
\begin{pmatrix} 2 & 5 & 7 \\ 0 & 1 & -6 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix} &= 2 \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + 1 \begin{pmatrix} 5 \\ 1 \\ 1 \end{pmatrix} - 3 \begin{pmatrix} 7 \\ -6 \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} -12 \\ 19 \\ -2 \end{pmatrix}.
\end{aligned}
$$

III. Products of matrices

We now know how to multiply a *matrix* A $(n \times k)$ *by a column* $\mathbf{C}$ $(k \times 1)$. Suppose we have the same matrix A but instead of a matrix consisting of a single column $\mathbf{C}$ we have a matrix B made up of several columns:

$$B = (\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, \ldots, \mathbf{C}_m)$$

where each column $\mathbf{C}_j$ has k entries—i.e. is a $k \times 1$ matrix. Then B has k rows and m columns—B is $k \times m$. To define $A \cdot B$, we simply multiply A by each of the columns $\mathbf{C}_j$ and put these columns $A \cdot \mathbf{C}_j$ side-by-side as columns of a new matrix:

$$\begin{aligned} A \cdot B &= A \cdot (\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, \ldots, \mathbf{C}_m) \\ &= (A \cdot \mathbf{C}_1, A \cdot \mathbf{C}_2, A \cdot \mathbf{C}_3, \ldots, A \cdot \mathbf{C}_m). \end{aligned}$$

$$\boxed{A \cdot B = \textbf{the matrix whose columns are } A \textbf{ times the columns of } B.}$$

Example 5.

$$\overbrace{\begin{pmatrix} 2 & 5 & 7 \\ 0 & 1 & -6 \\ 0 & 1 & 1 \end{pmatrix}}^{A} \cdot \overbrace{\begin{pmatrix} 2 & 1 & -2 \\ 1 & 3 & 0 \\ -3 & 2 & 4 \end{pmatrix}}^{B} = \begin{pmatrix} -12 & 31 & 24 \\ 19 & -9 & -24 \\ -2 & 5 & 4 \end{pmatrix}.$$

$$\boxed{A \cdot B = P = (p_{ij}) \text{ where } p_{ij} = (i\text{th row of } A) \cdot (j\text{th column of } B).}$$

There are several things worth noting. First of all, the dimensions of the matrices have to match up in order that their product is defined. Since $A \cdot B$ involves dotting the rows of A with the columns of B, we see that the rows of A have to have the same size as the columns of B. So, if A is $n \times k$ and B is $k' \times m$, then k must equal k'. Since the rows of $A \cdot B$ are obtained by dotting with the rows of A, $A \cdot B$ must have n rows. Since the columns of $A \cdot B$ are the products of A with the columns of B, $A \cdot B$ must have m columns. Schematically, this looks like:

$$\overbrace{(n \quad \times \quad \underbrace{k) \ \cdot \ (k'}_{k \text{ must equal } k'} \quad \times \quad m}^{\text{the product will be } n \times m}).$$

The other thing about matrix multiplication is that it is not commutative. This means that, in general, $A \cdot B$ will not be the same as $B \cdot A$.

Example 6. (See example 5).

$$\begin{aligned} \overbrace{\begin{pmatrix} 2 & 1 & -2 \\ 1 & 3 & 0 \\ -3 & 2 & 4 \end{pmatrix}}^{B} \cdot \overbrace{\begin{pmatrix} 2 & 5 & 7 \\ 0 & 1 & -6 \\ 0 & 1 & 1 \end{pmatrix}}^{A} &= \begin{pmatrix} 4 & 9 & 6 \\ 2 & 8 & -11 \\ -6 & -9 & -29 \end{pmatrix} \\ &\neq A \cdot B. \end{aligned}$$