Properties of Matrix Operations

In the previous section, we learned three operations on matrices: scalar multiplication, matrix addition, and matrix multiplication. These are indeed *new* operations to us, and so we need to discuss their properties in detail. To understand why we need to discuss their properties, consider the following example:

Example. Given

$$A = \begin{pmatrix} 2 & 0 & -1 \end{pmatrix}, \; B = \begin{pmatrix} 4 & 2 \\ -3 & 10 \\ 1 & -2 \end{pmatrix}, \text{ and } C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

compute ABC.

In order to compute ABC, we have to perform matrix multiplication *two times.* There's a bit of a problem here though–which product should we compute first? Potentially, we could:

1. multiply A and B to get the matrix AB, then calculate $(AB)C$, or
2. multiply B and C first to get the matrix BC, then calculate $A(BC)$.

This leads us to a problem–is repeated matrix multiplication "well-defined"? I.e., will we get the same answer regardless of how we go about performing the operations? As we have only just been introduced to the operation of matrix multiplication, we don't really have an answer to the question: $(AB)C$ might be the same matrix as $A(BC)$, or it might not. If they are different, then which one is the "right" way to make the calculation?

Fortunately for us, the list of properties below answers many such questions about the operations we learned in the previous section.

Properties of Scalar Multiplication, Matrix Addition, and Matrix Multiplication

Theorem. Assume that matrices A, B, and C have sizes that are amenable for performing the given operations, and let a and b be scalars. Then:

(a) $A + B = B + A$ (matrix addition is commutative)

(b) $(A + B) + C = A + (B + C)$ (matrix addition is associative)

(c) $(AB)C = A(BC)$ (matrix multiplication is associative)

(d) $A(B + C) = AB + AC$ and $(A + B)C = AC + BC$ (matrix multiplication distributes over addition)

(h) $a(B + C) = aB + aC$ (scalar multiplication distributes over addition)

(j) $(a + b)C = aC + bC$

(l) $a(bC) = (ab)C$

(m) $a(BC) = (aB)C = B(aC)$

Most of the properties above are not particularly exciting; indeed, most of them are exactly what we would expect given our understanding of scalar addition and scalar multiplication. Regardless, we point out a few features of some of the properties:

(b) $(A+B)+C = A+(B+C)$: If we wish to add three matrices, we have to decide which two of the three to add first; this property simply says that the order doesn't matter–we'll end up with the same matrix either way.

(c) $(AB)C = A(BC)$: Similarly, if we wish to *multiply* three matrices, we can start by multiplying AB, and then C, or we can multiply A by the (already computed) product BC; either way, we'll end up with the same matrix. This idea is illustrated in the following example.

(m) $a(BC) = (aB)C = B(aC)$: In a sense, scalar multiplication commutes with matrix multiplication–scalars can move through products at will without affecting the outcome.

We are going to prove part (c) of the theorem: $(AB)C = A(BC)$. To make life a bit simpler, we will only work through the case where A, B, and C are all square matrices, say $n \times n$. However, you should not that the theorem *always* works, as long as it is possible to multiply A, B, and C in that order; say for example if A is 3×4, B is 4×7, and C is 7×2.

Before we begin the proof, we should make a few notes about the proof technique that we will employ. We wish to show that two matrices–specifically $(AB)C$ and $A(BC)$–are the same. To do so, we need merely show that each entry of $(AB)C$ matches the corresponding entry of $A(BC)$, that is the i, j entry of the former must be the same as the i, j entry of the latter.

This fact can be established in general relatively quickly if we remember the summation notation that we used in the previous section to calculate the i, j entry of a product: the i, j entry of the product AB is the number

$$p_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Of course, since there are *two* products involved in calculating $(AB)C$, we will need to make the appropriate adjustments in our proof.

Theorem. If A, B, and C are $n \times n$ matrices, then

$$(AB)C = A(BC).$$

Proof. Let $A = [a_{ij}]$, $B = [b_{ij}]$, and $C = [c_{ij}]$, and $(AB)C = [q_{ij}]$. We begin by calculating the i, j entry q_{ij} of the product $(AB)C$.

Entry q_{ij} is the scalar product of row i of AB and column j of C. Of course, column j of C has form

$$\begin{pmatrix} c_{1j} \\ c_{2j} \\ \vdots \\ c_{kj} \\ \vdots \\ c_{nj} \end{pmatrix}.$$

We need to calculate the entries of row i of AB. The $i,1$ entry of AB is the scalar product of row i of A with column 1 of B, thus has form

$$\sum_{h=1}^{n} a_{ih} b_{h1}.$$

In general, the i,ℓ entry of the product AB is given by

$$\sum_{h=1}^{n} a_{ih} b_{h\ell}.$$

Thus row i of AB has form

$$\begin{pmatrix} \sum_{h=1}^{n} a_{ih} b_{h1} & \sum_{h=1}^{n} a_{ih} b_{h2} & \ldots & \sum_{h=1}^{n} a_{ih} b_{h\ell} & \ldots & \sum_{h=1}^{n} a_{ih} b_{hn} \end{pmatrix}.$$

Thus entry q_{ij} of $(AB)C$ is the number

$$\begin{aligned} q_{ij} &= \left(\sum_{h=1}^{n} a_{ih} b_{h1}\right) c_{1j} + \left(\sum_{h=1}^{n} a_{ih} b_{h2}\right) c_{2j} + \ldots + \left(\sum_{h=1}^{n} a_{ih} b_{hn}\right) c_{nj} \\ &= \sum_{\ell=1}^{n} \left(\left(\sum_{h=1}^{n} a_{ih} b_{h\ell} \right) c_{\ell j} \right). \end{aligned}$$

Similarly, we can calculate the i,j entry u_{ij} of $A(BC)$: row i of A has form

$$\begin{pmatrix} a_{i1} & a_{i2} & \ldots & a_{in} \end{pmatrix},$$

and column j of BC has form

$$\begin{pmatrix} \sum_{\ell=1}^{n} b_{1\ell} c_{\ell j} \\ \sum_{\ell=1}^{n} b_{2\ell} c_{\ell j} \\ \vdots \\ \sum_{\ell=1}^{n} b_{n\ell} c_{\ell j} \end{pmatrix}$$

$$u_{ij} = \sum_{h=1}^{n}\left(a_{ih}\sum_{\ell=1}^{n} b_{h\ell}c_{\ell j}\right).$$

Finally, we need to show that $q_{ij} = u_{ij}$. To do so, we expand and rewrite q_{ij}:

$$\begin{aligned}
q_{ij} &= \left(\sum_{h=1}^{n} a_{ih}b_{h1}\right)c_{1j} + \left(\sum_{h=1}^{n} a_{ih}b_{h2}\right)c_{2j} + \ldots + \left(\sum_{h=1}^{n} a_{ih}b_{hn}\right)c_{nj} \\
&= \left(a_{i1}b_{11}c_{1j} + \ldots + a_{in}b_{n1}c_{1j}\right) + \left(a_{i1}b_{12}c_{2j} + \ldots + a_{in}b_{n2}c_{2j}\right) \\
&\quad + \ldots + \left(a_{i1}b_{1n}c_{nj} + \ldots + a_{in}b_{nn}c_{nj}\right) \\
&= \left(a_{i1}b_{11}c_{1j} + a_{i1}b_{12}c_{2j} + \ldots + a_{i1}b_{1n}c_{nj}\right) + \ldots + \left(a_{in}b_{n1}c_{1j} + a_{in}b_{n2}c_{2j} + \ldots + a_{in}b_{nn}c_{nj}\right) \\
&= \left(a_{i1}\sum_{\ell=1}^{n} b_{1\ell}c_{\ell j}\right) + \ldots + \left(a_{in}\sum_{\ell=1}^{n} b_{n\ell}c_{\ell j}\right) \\
&= \sum_{h=1}^{n}\left(a_{ih}\sum_{\ell=1}^{n} b_{h\ell}c_{\ell j}\right) \\
&= u_{ij}.
\end{aligned}$$

Remark. While many of the rules in the theorem above look familiar from real number arithmetic, you may have noticed that one familiar rule is missing–the commutativity of multiplication. Indeed, we have already seen an example that indicates that matrix multiplication is, in general, *not* commutative:

$$\begin{pmatrix} 1 & 4i \\ 3 & 2 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 5 & -2 & 1 \\ 0 & -4 & 3i \end{pmatrix} = \begin{pmatrix} 5 & -2-16i & -11 \\ 15 & -14 & 3+6i \\ -5 & 2 & -1 \end{pmatrix}$$

while

$$\begin{pmatrix} 5 & -2 & 1 \\ 0 & -4 & 3i \end{pmatrix}\begin{pmatrix} 1 & 4i \\ 3 & 2 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -2 & -4+20i \\ -12-3i & -8 \end{pmatrix}.$$

While there *are* occasions when $AB = BA$, for the most part multiplication is not commutative, and you are generally safe assuming that

$$AB \neq BA.$$

Example. Given

$$A = \begin{pmatrix} 2 & 0 & -1 \end{pmatrix},\ B = \begin{pmatrix} 4 & 2 \\ -3 & 10 \\ 1 & -2 \end{pmatrix},\ \text{and } C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

verify that $(AB)C = A(BC)$.

Let's start by computing the matrix AB:

$$\begin{aligned} AB &= \begin{pmatrix} 2 & 0 & -1 \end{pmatrix} \begin{pmatrix} 4 & 2 \\ -3 & 10 \\ 1 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 8+0-1 & 4+0+2 \end{pmatrix} \\ &= \begin{pmatrix} 7 & 6 \end{pmatrix} \end{aligned}$$

Now that we know that

$$AB = \begin{pmatrix} 7 & 6 \end{pmatrix},$$

let's calculate $(AB)C$:

$$\begin{aligned} (AB)C &= \begin{pmatrix} 7 & 6 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0-6 & 7+0 \end{pmatrix} \\ &= \begin{pmatrix} -6 & 7 \end{pmatrix}. \end{aligned}$$

Thus we have calculated that

$$(AB)C = \begin{pmatrix} -6 & 7 \end{pmatrix}.$$

On the other hand, we could start by finding the matrix BC:

$$\begin{aligned} BC &= \begin{pmatrix} 4 & 2 \\ -3 & 10 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0-2 & 4+0 \\ 0-10 & -3+0 \\ 0+2 & 1+0 \end{pmatrix} \\ &= \begin{pmatrix} -2 & 4 \\ -10 & -3 \\ 2 & 1 \end{pmatrix}. \end{aligned}$$

Now we can find $A(BC)$:

$$\begin{aligned} A(BC) &= \begin{pmatrix} 2 & 0 & -1 \end{pmatrix} \begin{pmatrix} -2 & 4 \\ -10 & -3 \\ 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -4+0-2 & 8+0-1 \end{pmatrix} \\ &= \begin{pmatrix} -6 & 7 \end{pmatrix} \end{aligned}$$

Finally, we see that

$$(AB)C = \begin{pmatrix} -6 & 7 \end{pmatrix} = A(BC);$$

as promised by the theorem, the outcome is the same either way we make the computation.

Properties of the 0 Matrix

In the first section, we introduced the $\mathbf{0}$ matrix, all of whose entries are 0. While the $\mathbf{0}$ matrix is a different object from the *number* 0, it turns out that the $\mathbf{0}$ matrix behaves like the number 0 in many ways.

For example, we know that adding 0 to a number does not alter the number: $0 + a = a$, and we say that the number 0 acts as the *additive identity* for real numbers. Similarly, adding the $\mathbf{0}$ *matrix* to a matrix of the same size does not alter the matrix: $\mathbf{0} + A = A$, and we call $\mathbf{0}$ the *additive identity* for matrices.

The $\mathbf{0}$ matrix has several important properties that parallel the properties of the number 0; they are outlined below:

Theorem. Let A be an $m \times n$ matrix, and $\mathbf{0}$ be a zero matrix of size $m \times n$ (for (a), (c), and (e)) or size $k \times m$ (for (d)). Let c be any real number. Then:

(a) $\mathbf{0} + A = A + \mathbf{0} = A$

(c) $A + (-A) = \mathbf{0}$

(d) $\mathbf{0}A = \mathbf{0}$

(e) If $cA = \mathbf{0}$, then either $c = 0$ or $A = \mathbf{0}$.

Remark. As we discuss parallels between real number arithmetic and matrix arithmetic, it is important to point out times when such parallels fail. Our discussion of the $\mathbf{0}$ matrix leads directly to two such failures:

1. Cancelation: For real numbers, we know that if

   $$ab = ac \text{ and } a \neq 0, \text{ then } b = c.$$

   There is no such law for matrices; indeed it is perfectly possible to produce identical products from different matrices, as indicated by the following example. Let

   $$A = \begin{pmatrix} 0 & 0 \\ -4 & 0 \end{pmatrix}, \ B = \begin{pmatrix} 1 & 1 \\ 5 & 3 \end{pmatrix}, \text{ and } C = \begin{pmatrix} 1 & 1 \\ 2 & 7 \end{pmatrix}.$$

   A few quick calculations show that, even though $B \neq C$, $AB = AC$:

   $$\begin{pmatrix} 0 & 0 \\ -4 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 5 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -4 & -4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -4 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 7 \end{pmatrix}.$$

2. 0 Products: If a and b are real numbers so that $ab = 0$, then we are guaranteed that at least one of a or b is 0. This is far from the case for matrices; setting

$$A = \begin{pmatrix} 0 & 0 \\ -4 & 0 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 & 0 \\ 1 & 2 \end{pmatrix},$$

it is easy to show that

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

even though neither of A nor B is $\mathbf{0}$.

Properties of the Identity Matrix

We introduced another important matrix in section 1: the identity matrix I_n is the $n \times n$ (square) matrix whose main diagonal entries are all 1s, and all of whose other entries are 0s. For example, I_4 is given by

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Just as the $\mathbf{0}$ matrix parallels the number 0, I shares properties with the number 1 (which is referred to as the *multiplicative identity* for real numbers).

For example, the number 1 has the property that $1 \cdot a = a \cdot 1 = a$; in a sense it does not alter numbers multiplicatively. Similarly, I acts as a multiplicative identity for matrices:

Theorem. If A is an $m \times n$ matrix, then

$$I_m A = A \text{ and } AI_n = A.$$

Transposes and Matrix Operations

Recall that the transpose of a matrix A is the matrix $A^\top$ that we build by switching the rows and columns of A. For example, if

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & 5 \end{pmatrix}, \text{ then } A^\top = \begin{pmatrix} 1 & 3 \\ 2 & 4 \\ 0 & 5 \end{pmatrix}.$$

The operation of transposing interacts nicely with the operations of addition, scalar multiplication, and matrix multiplication, as indicated in the following theorem:

Theorem. Assume that A and B are matrices of the appropriate sizes to perform the indicated operations, and let k be any scalar. Then

(a) $(A^\top)^\top = A$

(b) $(A+B)^\top = A^\top + B^\top$

(d) $(kA)^\top = kA^\top$

(e) $(AB)^\top = B^\top A^\top$.