Matrix Arithmetic

In Section **??**, we used a special class of matrices, the augmented matrices, to assist us in solving systems of linear equations. In this section, we study matrices as mathematical objects of their own accord, temporarily divorced from systems of linear equations. To do so conveniently requires some more notation. When we write $A = \left[a_{ij}\right]_{m \times n}$, we mean A is an m by n matrix[1] and a_{ij} is the entry found in the ith row and jth column. Schematically, we have

$$A = \begin{array}{c} \xrightarrow{\text{j counts columns from left to right}} \\ \left[\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{array}\right] \end{array} \Big\downarrow \; \text{i counts rows from top to bottom}$$

With this new notation we can define what it means for two matrices to be equal.

Definition 1. Matrix Equality: Two matrices are said to be **equal** if they are the same size and their corresponding entries are equal. More specifically, if $A = \left[a_{ij}\right]_{m \times n}$ and $B = \left[b_{ij}\right]_{p \times r}$, we write $A = B$ provided

1. $m = p$ and $n = r$
2. $a_{ij} = b_{ij}$ for all $1 \leq i \leq m$ and all $1 \leq j \leq n$.

Essentially, two matrices are equal if they are the same size and they have the same numbers in the same spots. For example, the two 2×3 matrices below are, despite appearances, equal.

$$\left[\begin{array}{rrr} 0 & -2 & 9 \\ 25 & 117 & -3 \end{array}\right] = \left[\begin{array}{ccc} \ln(1) & \sqrt[3]{-8} & e^{2\ln(3)} \\ 125^{2/3} & 3^2 \cdot 13 & \log(0.001) \end{array}\right]$$

Now that we have an agreed upon understanding of what it means for two matrices to equal each other, we may begin defining arithmetic operations on matrices. Our first operation is addition.

Definition 2. Matrix Addition: Given two matrices of the same size, the matrix obtained by adding the corresponding entries of the two matrices is called the **sum** of the two matrices. More specifically, if $A = \left[a_{ij}\right]_{m \times n}$ and $B = \left[b_{ij}\right]_{m \times n}$, we define

$$A + B = \left[a_{ij}\right]_{m \times n} + \left[b_{ij}\right]_{m \times n} = \left[a_{ij} + b_{ij}\right]_{m \times n}$$

As an example, consider the sum below.

$$\left[\begin{array}{rr} 2 & 3 \\ 4 & -1 \\ 0 & -7 \end{array}\right] + \left[\begin{array}{rr} -1 & 4 \\ -5 & -3 \\ 8 & 1 \end{array}\right] = \left[\begin{array}{rr} 2 + (-1) & 3 + 4 \\ 4 + (-5) & (-1) + (-3) \\ 0 + 8 & (-7) + 1 \end{array}\right] = \left[\begin{array}{rr} 1 & 7 \\ -1 & -4 \\ 8 & -6 \end{array}\right]$$

It is worth the reader's time to think what would have happened had we reversed the order of the summands above. As we would expect, we arrive at the same answer. In general, $A + B = B + A$ for matrices A and B, provided they are the same size so that the sum is defined in the first place. This is the *commutative property* of matrix addition. To see why this is true in general, we appeal to the definition of matrix addition. Given $A = \left[a_{ij}\right]_{m \times n}$ and $B = \left[b_{ij}\right]_{m \times n}$,

$$A + B = \left[a_{ij}\right]_{m \times n} + \left[b_{ij}\right]_{m \times n} = \left[a_{ij} + b_{ij}\right]_{m \times n} = \left[b_{ij} + a_{ij}\right]_{m \times n} = \left[b_{ij}\right]_{m \times n} + \left[a_{ij}\right]_{m \times n} = B + A$$

[1] Recall that means A has m rows and n columns.

where the second equality is the definition of $A + B$, the third equality holds by the commutative law of real number addition, and the fourth equality is the definition of $B+A$. In other words, matrix addition is commutative because real number addition is.

A similar argument shows the *associative property* of matrix addition also holds, inherited in turn from the associative law of real number addition. Specifically, for matrices A, B, and C of the same size, $(A + B) + C = A + (B + C)$. In other words, when adding more than two matrices, it doesn't matter how they are grouped. This means that we can write $A + B + C$ without parentheses and there is no ambiguity as to what this means.[2] These properties and more are summarized in the following theorem.

Theorem 1. Properties of Matrix Addition

- **Commutative Property:** For all $m \times n$ matrices, $A + B = B + A$
- **Associative Property:** For all $m \times n$ matrices, $(A + B) + C = A + (B + C)$
- **Additive Identity:** If $0_{m\times n}$ is the $m \times n$ matrix whose entries are all 0, for all $m \times n$ matrices A

$$A + 0_{m\times n} = 0_{m\times n} + A = A$$

  That is, the additive identity for a matrix is the matrix of the additive identity for each of its entries.

- **Additive Inverse:** For every given $m \times n$ matrix $A = \left[a_{ij}\right]_{m\times n}$, the matrix $B = \left[-a_{ij}\right]_{m\times n}$ satisfies

$$A + B = B + A = 0_{m\times n}$$

  That is, the additive inverse of a matrix is the matrix of the additive inverses of each of its entries.

The identity property is easily verified by resorting to the definition of matrix addition; just as the number 0 is the additive identity for real numbers, the matrix comprised of all 0's does the same job for matrices.

To establish the inverse property, we note that per the definition of matrix addition,

$$A + B = \left[a_{ij}\right]_{m\times n} + \left[-a_{ij}\right]_{m\times n} = \left[a_{ij} - a_{ij}\right]_{m\times n} = [0]_{m\times n} = 0_{m\times n}.$$

The fact that $B + A = 0_{m\times n}$ as well comes from the commutative property of matrix addition.

More about the additive inverse is true. If a matrix $C = \left[c_{ij}\right]_{m\times n}$ satisfies $A + C = 0_{m\times n}$, then once again by the definition of matrix addition, we must have $a_{ij} + c_{ij} = 0$, or $c_{ij} = -a_{ij}$ for all i and j. This shows the matrix C must be the matrix B as described in Theorem **??** which shows the additive inverse of a matrix is *unique*. In general, we denote the additive inverse of a matrix A using the (suggestive) symbol $-A$.

With the concept of additive inverse well in hand, we may now discuss what is meant by subtracting matrices. You may remember from arithmetic that $a - b = a + (-b)$; that is, subtraction is defined as 'adding the opposite (inverse).' We extend this concept to matrices. For two matrices A and B of the same size, we define $A - B = A + (-B)$. At the level of entries, this amounts to

$$A - B = A + (-B) = \left[a_{ij}\right]_{m\times n} + \left[-b_{ij}\right]_{m\times n} = \left[a_{ij} + \left(-b_{ij}\right)\right]_{m\times n} = \left[a_{ij} - b_{ij}\right]_{m\times n}$$

Thus to subtract two matrices of equal size, we subtract their corresponding entries. Surprised?

Our next task is to define what it means to multiply a matrix by a real number. Thinking back to arithmetic, you may recall that multiplication, at least by a natural number, can be thought of as 'rapid addition.' For example, $2 + 2 + 2 = 3 \cdot 2$. We know from algebra[3] that $3x = x + x + x$, so it seems natural that given a matrix A, we define $3A = A + A + A$. If $A = \left[a_{ij}\right]_{m\times n}$, we have

$$3A = A + A + A = \left[a_{ij}\right]_{m\times n} + \left[a_{ij}\right]_{m\times n} + \left[a_{ij}\right]_{m\times n} = \left[a_{ij} + a_{ij} + a_{ij}\right]_{m\times n} = \left[3a_{ij}\right]_{m\times n}$$

[2] We have seen this idea before in Sections **??** and **??**.

[3] The Distributive Property, in particular.

In other words, multiplying the *matrix* in this fashion by 3 is the same as multiplying *each entry* by 3. This leads us to the following definition.

Definition 3. Scalar[4] Multiplication: We define the product of a real number and a matrix to be the matrix obtained by multiplying each of its entries by said real number. More specifically, if k is a real number and $A = \left[a_{ij}\right]_{m\times n}$, we define

$$kA = k\left[a_{ij}\right]_{m\times n} = \left[ka_{ij}\right]_{m\times n}$$

The word 'scalar' means 'scaling factor' as we explain below. Every point $P(x, y)$ in the plane can be represented by its position matrix, P:

$$(x, y) \leftrightarrow P = \begin{bmatrix} x \\ y \end{bmatrix}$$

Suppose we take the point $(-2, 1)$ and multiply its position matrix by 3. We have

$$3P = 3\begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3(-2) \\ 3(1) \end{bmatrix} = \begin{bmatrix} -6 \\ 3 \end{bmatrix}.$$

This new matrix corresponds to the point $(-6, 3)$ which is the result of scaling both the horizontal and vertical directions by a factor of 3.

As did matrix addition, scalar multiplication inherits many properties from real number arithmetic. Below we summarize these properties.

Theorem 2. Properties of Scalar Multiplication

- **Associative Property:** For every $m \times n$ matrix A and scalars k and r, $(kr)A = k(rA)$.
- **Identity Property:** For all $m \times n$ matrices A, $1A = A$.
- **Additive Inverse Property:** For all $m \times n$ matrices A, $-A = (-1)A$.
- **Distributive Property of Scalar Multiplication over Scalar Addition:**
  For every $m \times n$ matrix A and scalars k and r,
  $$(k + r)A = kA + rA$$
- **Distributive Property of Scalar Multiplication over Matrix Addition:**
  For all $m \times n$ matrices A and B scalars k,
  $$k(A + B) = kA + kB$$
- **Zero Product Property:** If A is an $m \times n$ matrix and k is a scalar, then
  $$kA = 0_{m\times n} \quad \text{if and only if} \quad k = 0 \quad \text{or} \quad A = 0_{m\times n}$$

As with the other results in this section, Theorem **??** can be proved using the definitions of scalar multiplication and matrix addition. For example, to prove that $k(A + B) = kA + kB$ for a scalar k and $m \times n$ matrices A and B, we start by adding A and B, then multiplying by k and seeing how that compares with the sum of kA and kB.

$$k(A + B) = k\left(\left[a_{ij}\right]_{m\times n} + \left[b_{ij}\right]_{m\times n}\right) = k\left[a_{ij} + b_{ij}\right]_{m\times n} = \left[k\left(a_{ij} + b_{ij}\right)\right]_{m\times n} = \left[ka_{ij} + kb_{ij}\right]_{m\times n}$$

[4]The word 'scalar' here refers to real numbers. 'Scalar multiplication' in this context means we are multiplying a matrix by a real number (a scalar). We will discuss this term momentarily.

As for $kA + kB$, we have

$$kA + kB = k\left[a_{ij}\right]_{m\times n} + k\left[b_{ij}\right]_{m\times n} = \left[ka_{ij}\right]_{m\times n} + \left[kb_{ij}\right]_{m\times n} = \left[ka_{ij} + kb_{ij}\right]_{m\times n} \checkmark$$

which establishes the property. The remaining proofs are similar and are left to the reader.

The properties in Theorems **??** and **??** establish an algebraic system that lets us treat matrices and scalars more or less as we would real numbers and variables. In the following example, we challenge the reader to justify each and every step of the calculations using either properties of matrix arithmetic.

Example 1. Solve for the matrix A:

$$3A - \left(\begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix} + 5A\right) = \begin{bmatrix} -4 & 2 \\ 6 & -2 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 9 & 12 \\ -3 & 39 \end{bmatrix}$$

Solution.

$$\begin{aligned}
3A - \left(\begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix} + 5A\right) &= \begin{bmatrix} -4 & 2 \\ 6 & -2 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 9 & 12 \\ -3 & 39 \end{bmatrix} \\
3A + \left\{-\left(\begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix} + 5A\right)\right\} &= \begin{bmatrix} -4 & 2 \\ 6 & -2 \end{bmatrix} + \begin{bmatrix} \left(\frac{1}{3}\right)(9) & \left(\frac{1}{3}\right)(12) \\ \left(\frac{1}{3}\right)(-3) & \left(\frac{1}{3}\right)(39) \end{bmatrix} \\
3A + (-1)\left(\begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix} + 5A\right) &= \begin{bmatrix} -4 & 2 \\ 6 & -2 \end{bmatrix} + \begin{bmatrix} 3 & 4 \\ -1 & 13 \end{bmatrix} \\
3A + \left\{(-1)\begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix} + (-1)(5A)\right\} &= \begin{bmatrix} -1 & 6 \\ 5 & 11 \end{bmatrix} \\
3A + (-1)\begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix} + (-1)(5A) &= \begin{bmatrix} -1 & 6 \\ 5 & 11 \end{bmatrix} \\
3A + \begin{bmatrix} (-1)(2) & (-1)(-1) \\ (-1)(3) & (-1)(5) \end{bmatrix} + ((-1)(5))A &= \begin{bmatrix} -1 & 6 \\ 5 & 11 \end{bmatrix} \\
3A + \begin{bmatrix} -2 & 1 \\ -3 & -5 \end{bmatrix} + (-5)A &= \begin{bmatrix} -1 & 6 \\ 5 & 11 \end{bmatrix} \\
3A + (-5)A + \begin{bmatrix} -2 & 1 \\ -3 & -5 \end{bmatrix} &= \begin{bmatrix} -1 & 6 \\ 5 & 11 \end{bmatrix} \\
(3 + (-5))A + \begin{bmatrix} -2 & 1 \\ -3 & -5 \end{bmatrix} + \left(-\begin{bmatrix} -2 & 1 \\ -3 & -5 \end{bmatrix}\right) &= \begin{bmatrix} -1 & 6 \\ 5 & 11 \end{bmatrix} + \left(-\begin{bmatrix} -2 & 1 \\ -3 & -5 \end{bmatrix}\right) \\
(-2)A + 0_{2\times 2} &= \begin{bmatrix} -1 & 6 \\ 5 & 11 \end{bmatrix} - \begin{bmatrix} -2 & 1 \\ -3 & -5 \end{bmatrix} \\
(-2)A &= \begin{bmatrix} -1-(-2) & 6-1 \\ 5-(-3) & 11-(-5) \end{bmatrix} \\
(-2)A &= \begin{bmatrix} 1 & 5 \\ 8 & 16 \end{bmatrix} \\
\left(-\frac{1}{2}\right)((-2)A) &= -\frac{1}{2}\begin{bmatrix} 1 & 5 \\ 8 & 16 \end{bmatrix} \\
\left(\left(-\frac{1}{2}\right)(-2)\right)A &= \begin{bmatrix} \left(-\frac{1}{2}\right)(1) & \left(-\frac{1}{2}\right)(5) \\ \left(-\frac{1}{2}\right)(8) & \left(-\frac{1}{2}\right)(16) \end{bmatrix} \\
1A &= \begin{bmatrix} -\frac{1}{2} & -\frac{5}{2} \\ -4 & -\frac{16}{2} \end{bmatrix} \\
A &= \begin{bmatrix} -\frac{1}{2} & -\frac{5}{2} \\ -4 & -8 \end{bmatrix}
\end{aligned}$$

The reader is encouraged to check our answer in the original equation. ■

While the solution to the previous example is written in excruciating detail, in practice many of the steps above are omitted. The reader is encouraged to solve the equation in Example **??** as they would any other linear equation, for example: $3a - (2 + 5a) = -4 + \frac{1}{3}(9)$.

We now turn our attention to *matrix multiplication* - that is, multiplying a matrix by another matrix. Based on the 'no surprises' trend so far in the section, you may expect that in order to multiply two matrices, they must be of the same size and you find the product by multiplying the corresponding entries. While this kind of product is used in other areas of mathematics,[5] we define matrix multiplication to serve us in solving systems of linear equations.

To that end, we begin by defining the product of a row and a column. We motivate the general definition with an example. Consider the two matrices A and B below.

[5] See this article on the Hadamard Product.

$$A = \begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 1 & 2 & -8 \\ 4 & 8 & -5 & 9 \\ 5 & 0 & -2 & -12 \end{bmatrix}$$

Let $R1$ denote the first row of A and $C1$ denote the first column of B. To find the 'product' of $R1$ with $C1$, denoted $R1 \cdot C1$, we first find the product of the first entry in $R1$ and the first entry in $C1$. Next, we add to that the product of the second entry in $R1$ and the second entry in $C1$, and so on until we reach the last entry in $R1$ and the last entry in $C1$.

Using entry notation, $R1 \cdot C1 = a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} = (2)(3) + (0)(4) + (-1)(5) = 6 + 0 + (-5) = 1$. We can visualize this schematically as follows

$$\begin{bmatrix} \boxed{2} & \boxed{0} & \boxed{-1} \\ -10 & 3 & 5 \end{bmatrix} \begin{bmatrix} \boxed{3} & 1 & 2 & -8 \\ \boxed{4} & 8 & -5 & 9 \\ \boxed{5} & 0 & -2 & -12 \end{bmatrix}$$

$$\underbrace{\begin{matrix} \xrightarrow{\hspace{4em}} & \boxed{3} \\ \boxed{2} \quad 0 \quad -1 & 4 \\ & 5 \end{matrix}\Big\downarrow}_{\substack{a_{11}b_{11} \\ (2)(3)}} \; + \; \underbrace{\begin{matrix} \xrightarrow{\hspace{4em}} & 3 \\ 2 \quad \boxed{0} \quad -1 & \boxed{4} \\ & 5 \end{matrix}\Big\downarrow}_{\substack{a_{12}b_{21} \\ (0)(4)}} \; + \; \underbrace{\begin{matrix} \xrightarrow{\hspace{4em}} & 3 \\ 2 \quad 0 \quad \boxed{-1} & 4 \\ & \boxed{5} \end{matrix}\Big\downarrow}_{\substack{a_{13}b_{31} \\ (-1)(5)}}$$

To find $R2 \cdot C3$ where $R2$ denotes the second row of A and $C3$ denotes the third column of B, we proceed similarly. We start with finding the product of the first entry of $R2$ with the first entry in $C3$ then add to it the product of the second entry in $R2$ with the second entry in $C3$, and so forth. Using entry notation, we have $R2 \cdot C3 = a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} = (-10)(2) + (3)(-5) + (5)(-2) = -45$. Schematically,

$$\begin{bmatrix} 2 & 0 & -1 \\ \boxed{-10} & \boxed{3} & \boxed{5} \end{bmatrix} \begin{bmatrix} 3 & 1 & \boxed{2} & -8 \\ 4 & 8 & \boxed{-5} & 9 \\ 5 & 0 & \boxed{-2} & -12 \end{bmatrix}$$

$$\underbrace{\begin{matrix} \xrightarrow{\hspace{4em}} & \boxed{2} \\ \boxed{-10} \quad 3 \quad 5 & -5 \\ & -2 \end{matrix}\Big\downarrow}_{a_{21}b_{13} = (-10)(2) = -20} \; + \; \underbrace{\begin{matrix} \xrightarrow{\hspace{4em}} & 2 \\ -10 \quad \boxed{3} \quad 5 & \boxed{-5} \\ & -2 \end{matrix}\Big\downarrow}_{a_{22}b_{23} = (3)(-5) = -15} \; + \; \underbrace{\begin{matrix} \xrightarrow{\hspace{4em}} & 2 \\ -10 \quad 3 \quad \boxed{5} & -5 \\ & \boxed{-2} \end{matrix}\Big\downarrow}_{a_{23}b_{33} = (5)(-2) = -10}$$

Generalizing this process, we have the following definition.

Definition 4. Product of a Row and a Column: Suppose $A = [a_{ij}]_{m \times n}$ and $B = [b_{ij}]_{n \times r}$. Let Ri denote the ith row of A and let Cj denote the jth column of B. The **product of R_i and C_j, denoted $R_i \cdot C_j$** is the real number defined by

$$Ri \cdot Cj = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots a_{in}b_{nj}$$

Note that in order to multiply a row by a column, the number of entries in the row must match the number of entries in the column. We are now in the position to define matrix multiplication.

Definition 5. Matrix Multiplication: Suppose $A = [a_{ij}]_{m \times n}$ and $B = [b_{ij}]_{n \times r}$. Let Ri denote the ith row of A and let Cj denote the jth column of B. The **product of A and B**, denoted AB, is the matrix

$$AB = [Ri \cdot Cj]_{m \times r}$$

that is

$$AB = \begin{bmatrix} R1 \cdot C1 & R1 \cdot C2 & \ldots & R1 \cdot Cr \\ R2 \cdot C1 & R2 \cdot C2 & \ldots & R2 \cdot Cr \\ \vdots & \vdots & & \vdots \\ Rm \cdot C1 & Rm \cdot C2 & \ldots & Rm \cdot Cr \end{bmatrix}$$

There are a number of subtleties in Definition **??** which warrant closer inspection. First and foremost, Definition **??** tells us that the ij-entry of a matrix product AB is the ith row of A times the jth column of B. In order for this to be defined, the number of entries in the rows of A must match the number of entries in the columns of B. This means that the number of columns of A must match[6] the number of rows of B. In other words, to multiply A times B, the second dimension of A must match the first dimension of B, which is why in Definition **??**, $A_{m\times \underline{n}}$ is being multiplied by a matrix $B_{\underline{n}\times r}$.

Furthermore, the product matrix AB has as many rows as A and as many columns of B. As a result, when multiplying a matrix $A_{\underline{m}\times n}$ by a matrix $B_{n\times \underline{r}}$, the result is the matrix $AB_{\underline{m}\times \underline{r}}$.

Returning to our example matrices below, we see that A is a $2 \times \underline{3}$ matrix and B is a $\underline{3} \times 4$ matrix. This means that the product matrix AB is defined and will be a 2×4 matrix.

$$A = \begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 1 & 2 & -8 \\ 4 & 8 & -5 & 9 \\ 5 & 0 & -2 & -12 \end{bmatrix}$$

Using Ri to denote the ith row of A and Cj to denote the jth column of B, we form AB per to Definition **??**:

$$AB = \begin{bmatrix} R1 \cdot C1 & R1 \cdot C2 & R1 \cdot C3 & R1 \cdot C4 \\ R2 \cdot C1 & R2 \cdot C2 & R2 \cdot C3 & R2 \cdot C4 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 6 & -4 \\ 7 & 14 & -45 & 47 \end{bmatrix}$$

Note that the product BA is not defined, since B is a $3 \times \underline{4}$ matrix while A is a $\underline{2} \times 3$ matrix; B has more columns than A has rows, and so it is not possible to multiply a row of B by a column of A.

Even when the dimensions of A and B are compatible such that AB and BA are both defined, the product AB and BA aren't necessarily equal.[7] In other words, AB may not equal BA which means matrix multiplication is not, in general, commutative. That being said, several other real number properties are inherited by matrix multiplication, as illustrated in our next theorem.

Theorem 3. Properties of Matrix Multiplication Let A, B and C be matrices such that all of the matrix products below are defined and let k be a real number.

- **Associative Property of Matrix Multiplication:** $(AB)C = A(BC)$
- **Associative Property with Scalar Multiplication:** $k(AB) = (kA)B = A(kB)$
- **Identity Property:**

  For a natural number k, the $\boldsymbol{k \times k}$ **identity matrix**, denoted I_k, is defined by $I_k = \left[d_{ij}\right]_{k\times k}$ where

  $$d_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}$$

  For all $m \times n$ matrices, $I_m A = A I_n = A$.

[6]The reader is encouraged to think this through carefully.

[7]And may not even have the same dimensions. For example, if A is a 2×3 matrix and B is a 3×2 matrix, then AB is defined and is a 2×2 matrix while BA is also defined... but is a 3×3 matrix!

- **Distributive Property of Matrix Multiplication over Matrix Addition:**

$$A(B \pm C) = AB \pm AC \text{ and } (A \pm B)C = AC \pm BC$$

The one property in Theorem **??** which begs further investigation is, without doubt, the multiplicative identity. The entries in a matrix where $i = j$ comprise what is called the *main diagonal* of the matrix. The identity matrix has 1's along its main diagonal and 0's everywhere else. A few examples of the matrix I_k mentioned in Theorem **??** are given below. The reader is encouraged to see how they match the definition of the identity matrix presented there.

$$\underset{I_1}{[1]} \quad \underset{I_2}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} \quad \underset{I_3}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \quad \underset{I_4}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}$$

The identity matrix is an example of what is called a *square matrix* as it has the same number of rows as columns. Note that to in order to verify that the identity matrix acts as a multiplicative identity, some care must be taken depending on the order of the multiplication. For example, take the matrix 2×3 matrix A:

$$A = \begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix}$$

In order for the product I_kA to be defined, $k = 2$; similarly, for AI_k to be defined, $k = 3$. We leave it to the reader to show $I_2A = A$ and $AI_3 = A$. In other words,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix}$$

and

$$\begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix}$$

While the proofs of the properties in Theorem **??** are computational in nature, the notation becomes quite involved very quickly, so they are left to a course in Linear Algebra. The following example provides some practice with matrix multiplication and its properties. As usual, some valuable lessons are to be learned.

Example 2.

1. Find AB for $A = \begin{bmatrix} -23 & -1 & 17 \\ 46 & 2 & -34 \end{bmatrix}$ and $B = \begin{bmatrix} -3 & 2 \\ 1 & 5 \\ -4 & 3 \end{bmatrix}$.
2. Find $C^2 - 5C + 10I_2$ for $C = \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix}$.
3. Suppose M is a 4×4 matrix. Use Theorem **??** to expand $(M - 2I_4)(M + 3I_4)$.

Solution.

1. We have $AB = \begin{bmatrix} -23 & -1 & 17 \\ 46 & 2 & -34 \end{bmatrix} \begin{bmatrix} -3 & 2 \\ 1 & 5 \\ -4 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

2. Just as x^2 means x times itself, C^2 denotes the matrix C times itself. We get

$$\begin{aligned}
C^2 - 5C + 10I_2 &= \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix}^2 - 5\begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix} + 10\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} -5 & 10 \\ -15 & -20 \end{bmatrix} + \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} \\
&= \begin{bmatrix} -5 & -10 \\ 15 & 10 \end{bmatrix} + \begin{bmatrix} 5 & 10 \\ -15 & -10 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.
\end{aligned}$$

3. We expand $(M - 2I_4)(M + 3I_4)$ with the same pedantic zeal we showed in Example **??**. The reader is encouraged to determine which property of matrix arithmetic justifies each step.

$$\begin{aligned}
(M - 2I_4)(M + 3I_4) &= (M - 2I_4)M + (M - 2I_4)(3I_4) \\
&= MM - (2I_4)M + M(3I_4) - (2I_4)(3I_4) \\
&= M^2 - 2(I_4M) + 3(MI_4) - 2(I_4(3I_4)) \\
&= M^2 - 2M + 3M - 2(3(I_4I_4)) \\
&= M^2 + M - 6I_4
\end{aligned}$$

■

Example **??** illustrates some interesting features of matrix multiplication. First note that in the first problem, neither A nor B is the zero matrix, yet the product AB is the zero matrix. Hence, the the zero product property enjoyed by real numbers and scalar multiplication does not hold for matrix multiplication.

The second and third problems introduce us to polynomials involving matrices. The reader is encouraged to step back and compare our expansion of the matrix product $(M - 2I_4)(M + 3I_4)$ in third probem with the product $(x - 2)(x + 3)$ from real number algebra. The exercises explore this kind of parallel further.

As we mentioned earlier, a point $P(x, y)$ in the xy-plane can be represented as a 2×1 position matrix. We now show that matrix multiplication can be used to rotate these points, and hence graphs of equations.

Example 3. Let $R = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$.

1. Plot $P(2, -2)$, $Q(4, 0)$, $S(0, 3)$, and $T(-3, -3)$ in the plane as well as the points RP, RQ, RS, and RT. Plot the lines $y = x$ and $y = -x$ as guides. What does R appear to be doing to these points?

2. If a point P is on the hyperbola $x^2 - y^2 = 4$, show that the point RP is on the curve $y = \frac{2}{x}$.

Solution.

1. For $P(2, -2)$, the position matrix is $P = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$, and

$$\begin{aligned}
RP &= \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}\begin{bmatrix} 2 \\ -2 \end{bmatrix} \\
&= \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix}
\end{aligned}$$

We have that R takes $(2,-2)$ to $(2\sqrt{2},0)$. Similarly, we find $(4,0)$ is moved to $(2\sqrt{2},2\sqrt{2})$, $(0,3)$ is moved to $\left(-\frac{3\sqrt{2}}{2},\frac{3\sqrt{2}}{2}\right)$, and $(-3,-3)$ is moved to $(0,-3\sqrt{2})$. We plot these points below along with the lines $y=x$ and $y=-x$. We see that the matrix R is rotating these points counterclockwise by 45°.

#1

Plotting P and RP.

2. For a generic point $P(x,y)$ on the hyperbola $x^2-y^2=4$, we have

$$\begin{aligned}
RP &= \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\
&= \begin{bmatrix} \frac{\sqrt{2}}{2}x-\frac{\sqrt{2}}{2}y \\ \frac{\sqrt{2}}{2}x+\frac{\sqrt{2}}{2}y \end{bmatrix}
\end{aligned}$$

which means R takes (x,y) to $\left(\frac{\sqrt{2}}{2}x-\frac{\sqrt{2}}{2}y,\frac{\sqrt{2}}{2}x+\frac{\sqrt{2}}{2}y\right)$. To show that this point is on the curve $y=\frac{2}{x}$, we replace x with $\frac{\sqrt{2}}{2}x-\frac{\sqrt{2}}{2}y$ and y with $\frac{\sqrt{2}}{2}x+\frac{\sqrt{2}}{2}y$ and simplify.

$$\begin{aligned}
y &= \frac{2}{x} \\
\frac{\sqrt{2}}{2}x+\frac{\sqrt{2}}{2}y &\stackrel{?}{=} \frac{2}{\frac{\sqrt{2}}{2}x-\frac{\sqrt{2}}{2}y} \\
\left(\frac{\sqrt{2}}{2}x-\frac{\sqrt{2}}{2}y\right)\left(\frac{\sqrt{2}}{2}x+\frac{\sqrt{2}}{2}y\right) &\stackrel{?}{=} \left(\frac{2}{\frac{\sqrt{2}}{2}x-\frac{\sqrt{2}}{2}y}\right)\left(\frac{\sqrt{2}}{2}x-\frac{\sqrt{2}}{2}y\right) \\
\left(\frac{\sqrt{2}}{2}x\right)^2-\left(\frac{\sqrt{2}}{2}y\right)^2 &\stackrel{?}{=} 2 \\
\frac{x^2}{2}-\frac{y^2}{2} &\stackrel{?}{=} 2 \\
x^2-y^2 &\stackrel{\checkmark}{=} 4
\end{aligned}$$

Since (x,y) is on the hyperbola $x^2-y^2=4$, we know that this last equation is true. Since all of our steps are reversible, this last equation is equivalent to our original equation, showing the graph of $y=\frac{2}{x}$ is none other than the hyperbola $x^2-y^2=4$ when rotated counterclockwise by 45°. Below are the graphs of $x^2-y^2=4$ and $y=\frac{2}{x}$ for comparison.

#2

Graphing $x^2 - y^2 = 4$ and $\boldsymbol{y = \frac{2}{x}}$.

■

When we started this section, we mentioned that we would temporarily consider matrices as their own entities, but that the algebra developed here would ultimately allow us to solve systems of linear equations. To that end, consider the system

$$\left\{ \begin{array}{rcl} 3x - y + z & = & 8 \\ x + 2y - z & = & 4 \\ 2x + 3y - 4z & = & 10 \end{array} \right.$$

In Section **??**, we encoded this system into the augmented matrix

$$\left[\begin{array}{rrr|r} 3 & -1 & 1 & 8 \\ 1 & 2 & -1 & 4 \\ 2 & 3 & -4 & 10 \end{array} \right]$$

Recall that the entries to the left of the vertical line come from the coefficients of the variables in the system, while those on the right comprise the associated constants. For that reason, we may form the *coefficient matrix A*, the *unknowns matrix X* and the *constant matrix B* as below

$$A = \left[\begin{array}{rrr} 3 & -1 & 1 \\ 1 & 2 & -1 \\ 2 & 3 & -4 \end{array} \right] \quad X = \left[\begin{array}{c} x \\ y \\ z \end{array} \right] \quad B = \left[\begin{array}{r} 8 \\ 4 \\ 10 \end{array} \right]$$

We now consider the matrix equation $AX = B$.

$$\begin{array}{rcl} AX & = & B \\ \left[\begin{array}{rrr} 3 & -1 & 1 \\ 1 & 2 & -1 \\ 2 & 3 & -4 \end{array} \right] \left[\begin{array}{c} x \\ y \\ z \end{array} \right] & = & \left[\begin{array}{r} 8 \\ 4 \\ 10 \end{array} \right] \\ \left[\begin{array}{r} 3x - y + z \\ x + 2y - z \\ 2x + 3y - 4z \end{array} \right] & = & \left[\begin{array}{r} 8 \\ 4 \\ 10 \end{array} \right] \end{array}$$

We see that finding a solution (x, y, z) to the original system corresponds to finding a solution X for the matrix equation $AX = B$. If we think about solving the real number equation $ax = b$, we would simply 'divide' both sides by a. Is it possible to 'divide' both sides of the matrix equation $AX = B$ by the matrix A? This is the central topic of Section **??**.