

8.3 MATRIX ARITHMETIC

In Section 8.2, 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 $A = [a_{ij}]_{m \times n}$, we mean A is an m by n matrix¹ and a_{ij} is the entry found in the i th row and j th column. Schematically, we have

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

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

Definition 8.6. Matrix Equality: Two matrices are said to be **equal** if they are the same size and their corresponding entries are equal. More specifically, if $A = [a_{ij}]_{m \times n}$ and $B = [b_{ij}]_{p \times r}$, 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.² For example, the two 2×3 matrices below are, despite appearances, equal.

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

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 8.7. 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 $A = [a_{ij}]_{m \times n}$ and $B = [b_{ij}]_{m \times n}$, we define

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

As an example, consider the sum below.

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

²Critics may well ask: Why not leave it at that? Why the need for all the notation in Definition 8.6? It is the authors' attempt to expose you to the wonderful world of mathematical precision.

$$\begin{bmatrix} 2 & 3 \\ 4 & -1 \\ 0 & -7 \end{bmatrix} + \begin{bmatrix} -1 & 4 \\ -5 & -3 \\ 8 & 1 \end{bmatrix} = \begin{bmatrix} 2+(-1) & 3+4 \\ 4+(-5) & (-1)+(-3) \\ 0+8 & (-7)+1 \end{bmatrix} = \begin{bmatrix} 1 & 7 \\ -1 & -4 \\ 8 & -6 \end{bmatrix}$$

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 $A = [a_{ij}]_{m \times n}$ and $B = [b_{ij}]_{m \times n}$,

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

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.³ These properties and more are summarized in the following theorem.

Theorem 8.3. Properties of Matrix Addition

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

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

- **Inverse Property:** For every given $m \times n$ matrix A , there is a unique matrix denoted $-A$ called the **additive inverse of A** such that

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

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, given a matrix $A = [a_{ij}]_{m \times n}$, we are looking for a matrix $B = [b_{ij}]_{m \times n}$ so that $A + B = 0_{m \times n}$. By the definition of matrix addition, we must

³A technical detail which is sadly lost on most readers.

have that $a_{ij} + b_{ij} = 0$ for all i and j . Solving, we get $b_{ij} = -a_{ij}$. Hence, given a matrix A , its additive inverse, which we call $-A$, does exist and is unique and, moreover, is given by the formula: $-A = [-a_{ij}]_{m \times n}$. The long and short of this is: to get the additive inverse of a matrix, take additive inverses of each of its entries. 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

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

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 algebra⁴ that $3x = x + x + x$, so it seems natural that given a matrix A , we define $3A = A + A + A$. If $A = [a_{ij}]_{m \times n}$, we have

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

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 8.8. Scalar^a 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 $A = [a_{ij}]_{m \times n}$, we define

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

^aThe word ‘scalar’ here refers to real numbers. ‘Scalar multiplication’ in this context means we are multiplying a matrix by a real number (a scalar).

One may well wonder why the word ‘scalar’ is used for ‘real number.’ It has everything to do with ‘scaling’ factors.⁵ A 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

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

which corresponds to the point $(-6, 3)$. We can imagine taking $(-2, 1)$ to $(-6, 3)$ in this fashion as a dilation by a factor of 3 in both the horizontal and vertical directions. Doing this to all points (x, y) in the plane, therefore, has the effect of magnifying (scaling) the plane by a factor of 3.

⁴The Distributive Property, in particular.

⁵See Section 1.7.

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

Theorem 8.4. 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 8.4 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 .

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

As for $kA + kB$, we have

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

which establishes the property. The remaining properties are left to the reader. The properties in Theorems 8.3 and 8.4 establish an algebraic system that lets us treat matrices and scalars more or less as we would real numbers and variables, as the next example illustrates.

Example 8.3.1. Solve for the matrix A : $3A - \left(\begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix} + 5A \right) = \begin{bmatrix} -4 & 2 \\ 6 & -2 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} 9 & 12 \\ -3 & 39 \end{bmatrix}$ using the definitions and properties of matrix arithmetic.

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} (\frac{1}{3})(9) & (\frac{1}{3})(12) \\ (\frac{1}{3})(-3) & (\frac{1}{3})(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} \\
 (-\frac{1}{2})((-2)A) &= -\frac{1}{2} \begin{bmatrix} 1 & 5 \\ 8 & 16 \end{bmatrix} \\
 ((-\frac{1}{2})(-2))A &= \begin{bmatrix} (-\frac{1}{2})(1) & (-\frac{1}{2})(5) \\ (-\frac{1}{2})(8) & (-\frac{1}{2})(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. We have spelled out each step in this example to encourage the reader to justify each step using the definitions and properties we have established thus far for matrix arithmetic. The reader is encouraged to solve the equation in Example 8.3.1 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,⁶ 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.

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

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 $C1$. Finally, we take that sum and we add to that the product of 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

$$\begin{array}{c} \begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix} \begin{bmatrix} 3 & 1 & 2 & -8 \\ 4 & 8 & -5 & 9 \\ 5 & 0 & -2 & -12 \end{bmatrix} \\ \hline \begin{array}{ccc} \begin{array}{c} \xrightarrow{\quad} \boxed{3} \\ \boxed{2} \quad 0 \quad -1 \\ \downarrow \\ 4 \\ 5 \end{array} & + & \begin{array}{c} \xrightarrow{\quad} 3 \\ 2 \quad \boxed{0} \quad -1 \\ \downarrow \\ \boxed{4} \\ 5 \end{array} & + & \begin{array}{c} \xrightarrow{\quad} \quad 3 \\ 2 \quad 0 \quad \boxed{-1} \\ \downarrow \\ 4 \\ \boxed{5} \end{array} \\ \hline a_{11}b_{11} & + & a_{12}b_{21} & + & a_{13}b_{31} \\ (2)(3) & + & (0)(4) & + & (-1)(5) \end{array} \end{array}$$

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,

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

⁶See this article on the [Hadamard Product](#).

Using R_i to denote the i th row of A and C_j to denote the j th column of B , we form AB according to Definition 8.10.

$$AB = \begin{bmatrix} R_1 \cdot C_1 & R_1 \cdot C_2 & R_1 \cdot C_3 & R_1 \cdot C_4 \\ R_2 \cdot C_1 & R_2 \cdot C_2 & R_2 \cdot C_3 & R_2 \cdot C_4 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 6 & -4 \\ 7 & 14 & -45 & 47 \end{bmatrix}$$

Note that the product BA is not defined, since B is a 3×4 matrix while A is a 2×3 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.⁸ In other words, AB may not equal BA . Although there is no commutative property of matrix multiplication in general, several other real number properties are inherited by matrix multiplication, as illustrated in our next theorem.

Theorem 8.5. 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 $k \times k$ **identity matrix**, denoted I_k , is defined by $I_k = [d_{ij}]_{k \times k}$ where

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

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

- **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 8.5 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 8.5 are given below. The reader is encouraged to see how they match the definition of the identity matrix presented there.

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

⁸And 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!

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 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 from earlier

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

In order for the product $I_k A$ to be defined, $k = 2$; similarly, for AI_k to be defined, $k = 3$. We leave it to the reader to show $I_2 A = A$ 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 8.5 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 8.3.2.

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 8.5 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 8.3.1. The reader is encouraged to determine which property of matrix arithmetic is used as we proceed from one step to the next.

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

□

Example 8.3.2 illustrates some interesting features of matrix multiplication. First note that in part 1, 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. Parts 2 and 3 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 part 3 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.

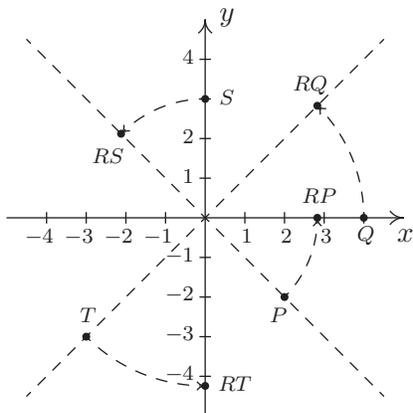
Example 8.3.3. Let $R = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$.

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. 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 $(-\frac{3\sqrt{2}}{2}, \frac{3\sqrt{2}}{2})$, and $(-3, -3)$ is moved to $(0, -3\sqrt{2})$. Plotting these in the coordinate plane along with the lines $y = x$ and $y = -x$, we see that the matrix R is rotating these points counterclockwise by 45° .



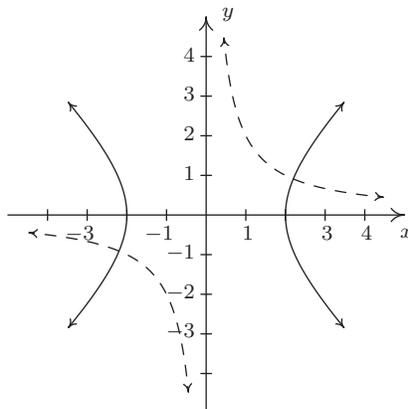
For a generic point $P(x, y)$ on the hyperbola $x^2 - y^2 = 4$, we have

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

which means R takes (x, y) to $(\frac{\sqrt{2}}{2}x - \frac{\sqrt{2}}{2}y, \frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}y)$. 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, which establishes the point is, indeed, on the graph of $y = \frac{2}{x}$. This means the graph of $y = \frac{2}{x}$ is a hyperbola, and it is none other than the hyperbola $x^2 - y^2 = 4$ rotated counterclockwise by 45° .⁹ Below we have the graph of $x^2 - y^2 = 4$ (solid line) and $y = \frac{2}{x}$ (dashed line) for comparison.



□

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

$$\begin{cases} 3x - y + z = 8 \\ x + 2y - z = 4 \\ 2x + 3y - 4z = 10 \end{cases}$$

In Section 8.2, we encoded this system into the augmented matrix

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

⁹See Section 7.5 for more details.

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

$$A = \begin{bmatrix} 3 & -1 & 1 \\ 1 & 2 & -1 \\ 2 & 3 & -4 \end{bmatrix} \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad B = \begin{bmatrix} 8 \\ 4 \\ 10 \end{bmatrix}$$

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

$$AX = B$$

$$\begin{bmatrix} 3 & -1 & 1 \\ 1 & 2 & -1 \\ 2 & 3 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \\ 10 \end{bmatrix}$$

$$\begin{bmatrix} 3x - y + z \\ x + 2y - z \\ 2x + 3y - 4z \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \\ 10 \end{bmatrix}$$

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 8.4.

In Exercises 23 - 29, consider the following scenario. In the small village of Pedimaxus in the country of Sasquatchia, all 150 residents get one of the two local newspapers. Market research has shown that in any given week, 90% of those who subscribe to the Pedimaxus Tribune want to keep getting it, but 10% want to switch to the Sasquatchia Picayune. Of those who receive the Picayune, 80% want to continue with it and 20% want switch to the Tribune. We can express this situation using matrices. Specifically, let X be the ‘state matrix’ given by

$$X = \begin{bmatrix} T \\ P \end{bmatrix}$$

where T is the number of people who get the Tribune and P is the number of people who get the Picayune in a given week. Let Q be the ‘transition matrix’ given by

$$Q = \begin{bmatrix} 0.90 & 0.20 \\ 0.10 & 0.80 \end{bmatrix}$$

such that QX will be the state matrix for the next week.

23. Let’s assume that when Pedimaxus was founded, all 150 residents got the Tribune. (Let’s call this Week 0.) This would mean

$$X = \begin{bmatrix} 150 \\ 0 \end{bmatrix}$$

Since 10% of that 150 want to switch to the Picayune, we should have that for Week 1, 135 people get the Tribune and 15 people get the Picayune. Show that QX in this situation is indeed

$$QX = \begin{bmatrix} 135 \\ 15 \end{bmatrix}$$

24. Assuming that the percentages stay the same, we can get to the subscription numbers for Week 2 by computing Q^2X . How many people get each paper in Week 2?
25. Explain why the transition matrix does what we want it to do.
26. If the conditions do not change from week to week, then Q remains the same and we have what’s known as a **Stochastic Process**¹⁰ because Week n ’s numbers are found by computing Q^nX . Choose a few values of n and, with the help of your classmates and calculator, find out how many people get each paper for that week. You should start to see a pattern as $n \rightarrow \infty$.
27. If you didn’t see the pattern, we’ll help you out. Let

$$X_s = \begin{bmatrix} 100 \\ 50 \end{bmatrix}.$$

Show that $QX_s = X_s$. This is called the **steady state** because the number of people who get each paper didn’t change for the next week. Show that $Q^nX \rightarrow X_s$ as $n \rightarrow \infty$.

¹⁰More specifically, we have a Markov Chain, which is a special type of stochastic process.

28. Now let

$$S = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

Show that $Q^n \rightarrow S$ as $n \rightarrow \infty$.

29. Show that $SY = X_s$ for any matrix Y of the form

$$Y = \begin{bmatrix} & y \\ 150 - y & \end{bmatrix}$$

This means that no matter how the distribution starts in Pedimaxus, if Q is applied often enough, we always end up with 100 people getting the Tribune and 50 people getting the Picayune.

30. Let $z = a + bi$ and $w = c + di$ be arbitrary complex numbers. Associate z and w with the matrices

$$Z = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \quad \text{and} \quad W = \begin{bmatrix} c & d \\ -d & c \end{bmatrix}$$

Show that complex number addition, subtraction and multiplication are mirrored by the associated *matrix* arithmetic. That is, show that $Z + W$, $Z - W$ and ZW produce matrices which can be associated with the complex numbers $z + w$, $z - w$ and zw , respectively.

31. Let

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & -3 \\ -5 & 2 \end{bmatrix}$$

Compare $(A + B)^2$ to $A^2 + 2AB + B^2$. Discuss with your classmates what constraints must be placed on two arbitrary matrices A and B so that both $(A + B)^2$ and $A^2 + 2AB + B^2$ exist. When will $(A + B)^2 = A^2 + 2AB + B^2$? In general, what is the correct formula for $(A + B)^2$?

In Exercises 32 - 36, consider the following definitions. A square matrix is said to be an **upper triangular matrix** if all of its entries below the main diagonal are zero and it is said to be a **lower triangular matrix** if all of its entries above the main diagonal are zero. For example,

$$E = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & -9 \\ 0 & 0 & -5 \end{bmatrix}$$

from Exercises 8 - 21 above is an upper triangular matrix whereas

$$F = \begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix}$$

is a lower triangular matrix. (Zeros are allowed on the main diagonal.) Discuss the following questions with your classmates.

32. Give an example of a matrix which is neither upper triangular nor lower triangular.
33. Is the product of two $n \times n$ upper triangular matrices always upper triangular?
34. Is the product of two $n \times n$ lower triangular matrices always lower triangular?
35. Given the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

write A as LU where L is a lower triangular matrix and U is an upper triangular matrix?

36. Are there any matrices which are simultaneously upper and lower triangular?

8.3.2 ANSWERS

1. For $A = \begin{bmatrix} 2 & -3 \\ 1 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & -2 \\ 4 & 8 \end{bmatrix}$

• $3A = \begin{bmatrix} 6 & -9 \\ 3 & 12 \end{bmatrix}$

• $-B = \begin{bmatrix} -5 & 2 \\ -4 & -8 \end{bmatrix}$

• $A^2 = \begin{bmatrix} 1 & -18 \\ 6 & 13 \end{bmatrix}$

• $A - 2B = \begin{bmatrix} -8 & 1 \\ -7 & -12 \end{bmatrix}$

• $AB = \begin{bmatrix} -2 & -28 \\ 21 & 30 \end{bmatrix}$

• $BA = \begin{bmatrix} 8 & -23 \\ 16 & 20 \end{bmatrix}$

2. For $A = \begin{bmatrix} -1 & 5 \\ -3 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 10 \\ -7 & 1 \end{bmatrix}$

• $3A = \begin{bmatrix} -3 & 15 \\ -9 & 18 \end{bmatrix}$

• $-B = \begin{bmatrix} -2 & -10 \\ 7 & -1 \end{bmatrix}$

• $A^2 = \begin{bmatrix} -14 & 25 \\ -15 & 21 \end{bmatrix}$

• $A - 2B = \begin{bmatrix} -5 & -15 \\ 11 & 4 \end{bmatrix}$

• $AB = \begin{bmatrix} -37 & -5 \\ -48 & -24 \end{bmatrix}$

• $BA = \begin{bmatrix} -32 & 70 \\ 4 & -29 \end{bmatrix}$

3. For $A = \begin{bmatrix} -1 & 3 \\ 5 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 7 & 0 & 8 \\ -3 & 1 & 4 \end{bmatrix}$

• $3A = \begin{bmatrix} -3 & 9 \\ 15 & 6 \end{bmatrix}$

• $-B = \begin{bmatrix} -7 & 0 & -8 \\ 3 & -1 & -4 \end{bmatrix}$

• $A^2 = \begin{bmatrix} 16 & 3 \\ 5 & 19 \end{bmatrix}$

• $A - 2B$ is not defined

• $AB = \begin{bmatrix} -16 & 3 & 4 \\ 29 & 2 & 48 \end{bmatrix}$

• BA is not defined

4. For $A = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & 3 & -5 \\ 7 & -9 & 11 \end{bmatrix}$

• $3A = \begin{bmatrix} 6 & 12 \\ 18 & 24 \end{bmatrix}$

• $-B = \begin{bmatrix} 1 & -3 & 5 \\ -7 & 9 & -11 \end{bmatrix}$

• $A^2 = \begin{bmatrix} 28 & 40 \\ 60 & 88 \end{bmatrix}$

• $A - 2B$ is not defined

• $AB = \begin{bmatrix} 26 & -30 & 34 \\ 50 & -54 & 58 \end{bmatrix}$

• BA is not defined

5. For $A = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$ and $B = [1 \ 2 \ 3]$

• $3A = \begin{bmatrix} 21 \\ 24 \\ 27 \end{bmatrix}$

• A^2 is not defined

• $AB = \begin{bmatrix} 7 & 14 & 21 \\ 8 & 16 & 24 \\ 9 & 18 & 27 \end{bmatrix}$

• $-B = [-1 \ -2 \ -3]$

• $A - 2B$ is not defined

• $BA = [50]$

6. For $A = \begin{bmatrix} 1 & -2 \\ -3 & 4 \\ 5 & -6 \end{bmatrix}$ and $B = [-5 \ 1 \ 8]$

• $3A = \begin{bmatrix} 3 & -6 \\ -9 & 12 \\ 15 & -18 \end{bmatrix}$

• A^2 is not defined

• AB is not defined

• $-B = [5 \ -1 \ -8]$

• $A - 2B$ is not defined

• $BA = [32 \ -34]$

7. For $A = \begin{bmatrix} 2 & -3 & 5 \\ 3 & 1 & -2 \\ -7 & 1 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 & 1 \\ 17 & 33 & 19 \\ 10 & 19 & 11 \end{bmatrix}$

• $3A = \begin{bmatrix} 6 & -9 & 15 \\ 9 & 3 & -6 \\ -21 & 3 & -3 \end{bmatrix}$

• $A^2 = \begin{bmatrix} -40 & -4 & 11 \\ 23 & -10 & 15 \\ -4 & 21 & -36 \end{bmatrix}$

• $AB = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

• $-B = \begin{bmatrix} -1 & -2 & -1 \\ -17 & -33 & -19 \\ -10 & -19 & -11 \end{bmatrix}$

• $A - 2B = \begin{bmatrix} 0 & -7 & 3 \\ -31 & -65 & -40 \\ -27 & -37 & -23 \end{bmatrix}$

• $BA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

8. $7B - 4A = \begin{bmatrix} -4 & -29 \\ -47 & -2 \end{bmatrix}$

9. $AB = \begin{bmatrix} -10 & 1 \\ -20 & -1 \end{bmatrix}$

$$10. BA = \begin{bmatrix} -9 & -12 \\ 1 & -2 \end{bmatrix}$$

$$12. ED = \begin{bmatrix} \frac{67}{3} & 11 \\ -\frac{178}{3} & -72 \\ -30 & -40 \end{bmatrix}$$

$$14. A - 4I_2 = \begin{bmatrix} -3 & 2 \\ 3 & 0 \end{bmatrix}$$

$$16. (A + B)(A - B) = \begin{bmatrix} -7 & 3 \\ 46 & 2 \end{bmatrix}$$

$$18. E^2 + 5E - 36I_3 = \begin{bmatrix} -30 & 20 & -15 \\ 0 & 0 & -36 \\ 0 & 0 & -36 \end{bmatrix}$$

20. CDE is undefined

$$22. E_1A = \begin{bmatrix} d & e & f \\ a & b & c \end{bmatrix} \quad E_1 \text{ interchanged } R1 \text{ and } R2 \text{ of } A.$$

$$E_2A = \begin{bmatrix} 5a & 5b & 5c \\ d & e & f \end{bmatrix} \quad E_2 \text{ multiplied } R1 \text{ of } A \text{ by } 5.$$

$$E_3A = \begin{bmatrix} a - 2d & b - 2e & c - 2f \\ d & e & f \end{bmatrix} \quad E_3 \text{ replaced } R1 \text{ in } A \text{ with } R1 - 2R2.$$

$$E_4 = \begin{bmatrix} 1 & 0 \\ 0 & -6 \end{bmatrix}$$

11. $E + D$ is undefined

$$13. CD + 2I_2A = \begin{bmatrix} \frac{238}{3} & -126 \\ \frac{863}{15} & \frac{361}{5} \end{bmatrix}$$

$$15. A^2 - B^2 = \begin{bmatrix} -8 & 16 \\ 25 & 3 \end{bmatrix}$$

$$17. A^2 - 5A - 2I_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$19. EDC = \begin{bmatrix} \frac{3449}{15} & -\frac{407}{6} & 99 \\ -\frac{9548}{15} & -\frac{101}{3} & -648 \\ -324 & -35 & -360 \end{bmatrix}$$

$$21. ABCEDI_2 = \begin{bmatrix} -\frac{90749}{15} & -\frac{28867}{5} \\ -\frac{156601}{15} & -\frac{47033}{5} \end{bmatrix}$$