2. Matrix Algebra

In the study of systems of linear equations in Chapter 1, we found it convenient to manipulate the augmented matrix of the system. Our aim was to reduce it to row-echelon form (using elementary row operations) and hence to write down all solutions to the system. In the present chapter we consider matrices for their own sake. While some of the motivation comes from linear equations, it turns out that matrices can be multiplied and added and so form an algebraic system somewhat analogous to the real numbers. This "matrix algebra" is useful in ways that are quite different from the study of linear equations. For example, the geometrical transformations obtained by rotating the euclidean plane about the origin can be viewed as multiplications by certain 2×2 matrices. These "matrix transformations" are an important tool in geometry and, in turn, the geometry provides a "picture" of the matrices. Furthermore, matrix algebra has many other applications, some of which will be explored in this chapter. This subject is quite old and was first studied systematically in 1858 by Arthur Cayley.[1]

2.1 Matrix Addition, Scalar Multiplication, and Transposition

A rectangular array of numbers is called a **matrix** (the plural is **matrices**), and the numbers are called the **entries** of the matrix. Matrices are usually denoted by uppercase letters: A, B, C, and so on. Hence,

$$A = \left[\begin{array}{rrr} 1 & 2 & -1 \\ 0 & 5 & 6 \end{array} \right] \quad B = \left[\begin{array}{rr} 1 & -1 \\ 0 & 2 \end{array} \right] \quad C = \left[\begin{array}{r} 1 \\ 3 \\ 2 \end{array} \right]$$

are matrices. Clearly matrices come in various shapes depending on the number of **rows** and **columns**. For example, the matrix A shown has 2 rows and 3 columns. In general, a matrix with m rows and n columns is referred to as an $\boldsymbol{m \times n}$ ***matrix*** or as having **size** $\boldsymbol{m \times n}$. Thus matrices A, B, and C above have sizes 2×3, 2×2, and 3×1, respectively. A matrix of size $1 \times n$ is called a **row matrix**, whereas one of size $m \times 1$ is called a **column matrix**. Matrices of size $n \times n$ for some n are called **square** matrices.

Each entry of a matrix is identified by the row and column in which it lies. The rows are numbered from the top down, and the columns are numbered from left to right. Then the $\boldsymbol{(i,\ j)}$**-entry** of a matrix is

[1] Arthur Cayley (1821-1895) showed his mathematical talent early and graduated from Cambridge in 1842 as senior wrangler. With no employment in mathematics in view, he took legal training and worked as a lawyer while continuing to do mathematics, publishing nearly 300 papers in fourteen years. Finally, in 1863, he accepted the Sadlerian professorship in Cambridge and remained there for the rest of his life, valued for his administrative and teaching skills as well as for his scholarship. His mathematical achievements were of the first rank. In addition to originating matrix theory and the theory of determinants, he did fundamental work in group theory, in higher-dimensional geometry, and in the theory of invariants. He was one of the most prolific mathematicians of all time and produced 966 papers.

the number lying simultaneously in row i and column j. For example,

$$\text{The } (1,\ 2)\text{-entry of } \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \text{ is } -1.$$

$$\text{The } (2,\ 3)\text{-entry of } \begin{bmatrix} 1 & 2 & -1 \\ 0 & 5 & 6 \end{bmatrix} \text{ is } 6.$$

A special notation is commonly used for the entries of a matrix. If A is an $m \times n$ matrix, and if the $(i,\ j)$-entry of A is denoted as a_{ij}, then A is displayed as follows:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}$$

This is usually denoted simply as $A = \left[a_{ij}\right]$. Thus a_{ij} is the entry in row i and column j of A. For example, a 3×4 matrix in this notation is written

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

It is worth pointing out a convention regarding rows and columns: *Rows are mentioned before columns.* For example:

- *If a matrix has size $m \times n$, it has m rows and n columns.*
- *If we speak of the $(i,\ j)$-entry of a matrix, it lies in row i and column j.*
- *If an entry is denoted a_{ij}, the first subscript i refers to the row and the second subscript j to the column in which a_{ij} lies.*

Two points $(x_1,\ y_1)$ and $(x_2,\ y_2)$ in the plane are equal if and only if[2] they have the same coordinates, that is $x_1 = x_2$ and $y_1 = y_2$. Similarly, two matrices A and B are called **equal** (written $A = B$) if and only if:

1. *They have the same size.*
2. *Corresponding entries are equal.*

If the entries of A and B are written in the form $A = \left[a_{ij}\right]$, $B = \left[b_{ij}\right]$, described earlier, then the second condition takes the following form:

$$A = \left[a_{ij}\right] = \left[b_{ij}\right] \text{ means } a_{ij} = b_{ij} \text{ for all } i \text{ and } j$$

[2]If p and q are statements, we say that p implies q if q is true whenever p is true. Then "p if and only if q" means that both p implies q and q implies p. See Appendix B for more on this.

Example 2.1.1

Given $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 0 & 1 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$ discuss the possibility that $A = B$, $B = C$, $A = C$.

Solution. $A = B$ is impossible because A and B are of different sizes: A is 2×2 whereas B is 2×3. Similarly, $B = C$ is impossible. But $A = C$ is possible provided that corresponding entries are equal: $\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$ means $a = 1$, $b = 0$, $c = -1$, and $d = 2$.

Matrix Addition

Definition 2.1 Matrix Addition

If A and B are matrices of the same size, their **sum** *$A + B$ is the matrix formed by adding corresponding entries.*

If $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ and $B = \begin{bmatrix} b_{ij} \end{bmatrix}$, this takes the form

$$A + B = \begin{bmatrix} a_{ij} + b_{ij} \end{bmatrix}$$

Note that addition is *not* defined for matrices of different sizes.

Example 2.1.2

If $A = \begin{bmatrix} 2 & 1 & 3 \\ -1 & 2 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & 6 \end{bmatrix}$, compute $A + B$.

Solution.

$$A + B = \begin{bmatrix} 2+1 & 1+1 & 3-1 \\ -1+2 & 2+0 & 0+6 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 2 \\ 1 & 2 & 6 \end{bmatrix}$$

Example 2.1.3

Find a, b, and c if $\begin{bmatrix} a & b & c \end{bmatrix} + \begin{bmatrix} c & a & b \end{bmatrix} = \begin{bmatrix} 3 & 2 & -1 \end{bmatrix}$.

Solution. Add the matrices on the left side to obtain

$$\begin{bmatrix} a+c & b+a & c+b \end{bmatrix} = \begin{bmatrix} 3 & 2 & -1 \end{bmatrix}$$

Because corresponding entries must be equal, this gives three equations: $a + c = 3$, $b + a = 2$, and $c + b = -1$. Solving these yields $a = 3$, $b = -1$, $c = 0$.

If A, B, and C are any matrices *of the same size*, then

$$A+B=B+A \qquad \text{(commutative law)}$$
$$A+(B+C)=(A+B)+C \qquad \text{(associative law)}$$

In fact, if $A=\left[a_{ij}\right]$ and $B=\left[b_{ij}\right]$, then the $(i,\ j)$-entries of $A+B$ and $B+A$ are, respectively, $a_{ij}+b_{ij}$ and $b_{ij}+a_{ij}$. Since these are equal for all i and j, we get

$$A+B=\left[\begin{array}{c} a_{ij}+b_{ij} \end{array}\right]=\left[\begin{array}{c} b_{ij}+a_{ij} \end{array}\right]=B+A$$

The associative law is verified similarly.

The $m\times n$ matrix in which every entry is zero is called the $m\times n$ **zero matrix** and is denoted as 0 (or 0_{mn} if it is important to emphasize the size). Hence,

$$0+X=X$$

holds for all $m\times n$ matrices X. The **negative** of an $m\times n$ matrix A (written $-A$) is defined to be the $m\times n$ matrix obtained by multiplying each entry of A by -1. If $A=\left[a_{ij}\right]$, this becomes $-A=\left[-a_{ij}\right]$. Hence,

$$A+(-A)=0$$

holds for all matrices A where, of course, 0 is the zero matrix of the same size as A.

A closely related notion is that of subtracting matrices. If A and B are two $m\times n$ matrices, their **difference** $A-B$ is defined by

$$A-B=A+(-B)$$

Note that if $A=\left[a_{ij}\right]$ and $B=\left[b_{ij}\right]$, then

$$A-B=\left[a_{ij}\right]+\left[-b_{ij}\right]=\left[a_{ij}-b_{ij}\right]$$

is the $m\times n$ matrix formed by *subtracting* corresponding entries.

Example 2.1.4

Let $A=\left[\begin{array}{rrr} 3 & -1 & 0 \\ 1 & 2 & -4 \end{array}\right]$, $B=\left[\begin{array}{rrr} 1 & -1 & 1 \\ -2 & 0 & 6 \end{array}\right]$, $C=\left[\begin{array}{rrr} 1 & 0 & -2 \\ 3 & 1 & 1 \end{array}\right]$. Compute $-A$, $A-B$, and $A+B-C$.

Solution.

$$-A=\left[\begin{array}{rrr} -3 & 1 & 0 \\ -1 & -2 & 4 \end{array}\right]$$

$$A-B=\left[\begin{array}{ccc} 3-1 & -1-(-1) & 0-1 \\ 1-(-2) & 2-0 & -4-6 \end{array}\right]=\left[\begin{array}{rrr} 2 & 0 & -1 \\ 3 & 2 & -10 \end{array}\right]$$

$$A+B-C=\left[\begin{array}{ccc} 3+1-1 & -1-1-0 & 0+1-(-2) \\ 1-2-3 & 2+0-1 & -4+6-1 \end{array}\right]=\left[\begin{array}{rrr} 3 & -2 & 3 \\ -4 & 1 & 1 \end{array}\right]$$

Example 2.1.5

Solve $\begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix} + X = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$ where X is a matrix.

Solution. We solve a numerical equation $a + x = b$ by subtracting the number a from both sides to obtain $x = b - a$. This also works for matrices. To solve $\begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix} + X = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$ simply subtract the matrix $\begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix}$ from both sides to get

$$X = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} - \begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1-3 & 0-2 \\ -1-(-1) & 2-1 \end{bmatrix} = \begin{bmatrix} -2 & -2 \\ 0 & 1 \end{bmatrix}$$

The reader should verify that this matrix X does indeed satisfy the original equation.

The solution in Example 2.1.5 solves the single matrix equation $A + X = B$ directly via matrix subtraction: $X = B - A$. This ability to work with matrices as entities lies at the heart of matrix algebra.

It is important to note that the sizes of matrices involved in some calculations are often determined by the context. For example, if

$$A + C = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 0 & 1 \end{bmatrix}$$

then A and C must be the same size (so that $A + C$ makes sense), and that size must be 2×3 (so that the sum is 2×3). For simplicity we shall often omit reference to such facts when they are clear from the context.

Scalar Multiplication

In gaussian elimination, multiplying a row of a matrix by a number k means multiplying *every* entry of that row by k.

Definition 2.2 Matrix Scalar Multiplication

*More generally, if A is any matrix and k is any number, the **scalar multiple** kA is the matrix obtained from A by multiplying each entry of A by k.*

If $A = \left[a_{ij} \right]$, this is

$$kA = \left[ka_{ij} \right]$$

Thus $1A = A$ and $(-1)A = -A$ for any matrix A.

The term *scalar* arises here because the set of numbers from which the entries are drawn is usually referred to as the set of scalars. We have been using real numbers as scalars, but we could equally well have been using complex numbers.

Example 2.1.6

If $A = \begin{bmatrix} 3 & -1 & 4 \\ 2 & 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 2 \end{bmatrix}$ compute $5A$, $\frac{1}{2}B$, and $3A - 2B$.

Solution.

$$5A = \begin{bmatrix} 15 & -5 & 20 \\ 10 & 0 & 30 \end{bmatrix}, \quad \tfrac{1}{2}B = \begin{bmatrix} \frac{1}{2} & 1 & -\frac{1}{2} \\ 0 & \frac{3}{2} & 1 \end{bmatrix}$$

$$3A - 2B = \begin{bmatrix} 9 & -3 & 12 \\ 6 & 0 & 18 \end{bmatrix} - \begin{bmatrix} 2 & 4 & -2 \\ 0 & 6 & 4 \end{bmatrix} = \begin{bmatrix} 7 & -7 & 14 \\ 6 & -6 & 14 \end{bmatrix}$$

If A is any matrix, note that kA is the same size as A for all scalars k. We also have

$$0A = 0 \quad \text{and} \quad k0 = 0$$

because the zero matrix has every entry zero. In other words, $kA = 0$ if either $k = 0$ or $A = 0$. The converse of this statement is also true, as Example 2.1.7 shows.

Example 2.1.7

If $kA = 0$, show that either $k = 0$ or $A = 0$.

Solution. Write $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ so that $kA = 0$ means $ka_{ij} = 0$ for all i and j. If $k = 0$, there is nothing to do. If $k \neq 0$, then $ka_{ij} = 0$ implies that $a_{ij} = 0$ for all i and j; that is, $A = 0$.

For future reference, the basic properties of matrix addition and scalar multiplication are listed in Theorem 2.1.1.

Theorem 2.1.1

Let A, B, and C denote arbitrary $m \times n$ matrices where m and n are fixed. Let k and p denote arbitrary real numbers. Then

1. $A + B = B + A$.
2. $A + (B + C) = (A + B) + C$.
3. *There is an $m \times n$ matrix 0, such that $0 + A = A$ for each A.*
4. *For each A there is an $m \times n$ matrix, $-A$, such that $A + (-A) = 0$.*
5. $k(A + B) = kA + kB$.
6. $(k + p)A = kA + pA$.
7. $(kp)A = k(pA)$.
8. $1A = A$.

Proof. Properties 1–4 were given previously. To check Property 5, let $A = \left[a_{ij}\right]$ and $B = \left[b_{ij}\right]$ denote matrices of the same size. Then $A + B = \left[a_{ij} + b_{ij}\right]$, as before, so the (i, j)-entry of $k(A+B)$ is

$$k(a_{ij} + b_{ij}) = ka_{ij} + kb_{ij}$$

But this is just the (i, j)-entry of $kA + kB$, and it follows that $k(A+B) = kA + kB$. The other Properties can be similarly verified; the details are left to the reader. □

The Properties in Theorem 2.1.1 enable us to do calculations with matrices in much the same way that numerical calculations are carried out. To begin, Property 2 implies that the sum

$$(A+B)+C = A+(B+C)$$

is the same no matter how it is formed and so is written as $A+B+C$. Similarly, the sum

$$A+B+C+D$$

is independent of how it is formed; for example, it equals both $(A+B)+(C+D)$ and $A+[B+(C+D)]$. Furthermore, property 1 ensures that, for example,

$$B+D+A+C = A+B+C+D$$

In other words, the *order* in which the matrices are added does not matter. A similar remark applies to sums of five (or more) matrices.

Properties 5 and 6 in Theorem 2.1.1 are called **distributive laws** for scalar multiplication, and they extend to sums of more than two terms. For example,

$$k(A+B-C) = kA+kB-kC$$

$$(k+p-m)A = kA+pA-mA$$

Similar observations hold for more than three summands. These facts, together with properties 7 and 8, enable us to simplify expressions by collecting like terms, expanding, and taking common factors in exactly the same way that algebraic expressions involving variables and real numbers are manipulated. The following example illustrates these techniques.

Example 2.1.8

Simplify $2(A+3C)-3(2C-B)-3\left[2(2A+B-4C)-4(A-2C)\right]$ where A, B, and C are all matrices of the same size.

Solution. The reduction proceeds as though A, B, and C were variables.

$$\begin{aligned} &2(A+3C)-3(2C-B)-3\left[2(2A+B-4C)-4(A-2C)\right] \\ &\quad = 2A+6C-6C+3B-3\left[4A+2B-8C-4A+8C\right] \\ &\quad = 2A+3B-3\left[2B\right] \\ &\quad = 2A-3B \end{aligned}$$

Transpose of a Matrix

Many results about a matrix A involve the *rows* of A, and the corresponding result for columns is derived in an analogous way, essentially by replacing the word *row* by the word *column* throughout. The following definition is made with such applications in mind.

Definition 2.3 Transpose of a Matrix

*If A is an $m \times n$ matrix, the **transpose** of A, written A^T, is the $n \times m$ matrix whose rows are just the columns of A in the same order.*

In other words, the first row of A^T is the first column of A (that is it consists of the entries of column 1 in order). Similarly the second row of A^T is the second column of A, and so on.

Example 2.1.9

Write down the transpose of each of the following matrices.

$$A = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \quad B = \begin{bmatrix} 5 & 2 & 6 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \quad D = \begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & 2 \\ -1 & 2 & 1 \end{bmatrix}$$

Solution.

$$A^T = \begin{bmatrix} 1 & 3 & 2 \end{bmatrix}, \ B^T = \begin{bmatrix} 5 \\ 2 \\ 6 \end{bmatrix}, \ C^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}, \text{ and } D^T = D.$$

If $A = [a_{ij}]$ is a matrix, write $A^T = [b_{ij}]$. Then b_{ij} is the jth element of the ith row of A^T and so is the jth element of the ith *column* of A. This means $b_{ij} = a_{ji}$, so the definition of A^T can be stated as follows:

$$\text{If } A = [a_{ij}], \text{ then } A^T = [a_{ji}]. \tag{2.1}$$

This is useful in verifying the following properties of transposition.

Theorem 2.1.2

Let A and B denote matrices of the same size, and let k denote a scalar.

1. *If A is an $m \times n$ matrix, then A^T is an $n \times m$ matrix.*
2. $(A^T)^T = A$.
3. $(kA)^T = kA^T$.
4. $(A+B)^T = A^T + B^T$.

Proof. Property 1 is part of the definition of A^T, and Property 2 follows from (2.1). As to Property 3: If $A = \left[a_{ij}\right]$, then $kA = \left[ka_{ij}\right]$, so (2.1) gives

$$(kA)^T = \left[ka_{ji}\right] = k\left[a_{ji}\right] = kA^T$$

Finally, if $B = \left[b_{ij}\right]$, then $A+B = \left[c_{ij}\right]$ where $c_{ij} = a_{ij} + b_{ij}$ Then (2.1) gives Property 4:

$$(A+B)^T = \left[c_{ij}\right]^T = \left[c_{ji}\right] = \left[a_{ji}+b_{ji}\right] = \left[a_{ji}\right] + \left[b_{ji}\right] = A^T + B^T$$

□

There is another useful way to think of transposition. If $A = \left[a_{ij}\right]$ is an $m \times n$ matrix, the elements $a_{11}, a_{22}, a_{33}, \dots$ are called the **main diagonal** of A. Hence the main diagonal extends down and to the right from the upper left corner of the matrix A; it is shaded in the following examples:

$$\left[\begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{array}\right] \left[\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{array}\right] \left[\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array}\right] \left[\begin{array}{c} a_{11} \\ a_{21} \end{array}\right]$$

Thus forming the transpose of a matrix A can be viewed as "flipping" A about its main diagonal, or as "rotating" A through 180° about the line containing the main diagonal. This makes Property 2 in Theorem 2.1.2 transparent.

Example 2.1.10

Solve for A if $\left(2A^T - 3\left[\begin{array}{rr} 1 & 2 \\ -1 & 1 \end{array}\right]\right)^T = \left[\begin{array}{rr} 2 & 3 \\ -1 & 2 \end{array}\right]$.

Solution. Using Theorem 2.1.2, the left side of the equation is

$$\left(2A^T - 3\left[\begin{array}{rr} 1 & 2 \\ -1 & 1 \end{array}\right]\right)^T = 2\left(A^T\right)^T - 3\left[\begin{array}{rr} 1 & 2 \\ -1 & 1 \end{array}\right]^T = 2A - 3\left[\begin{array}{rr} 1 & -1 \\ 2 & 1 \end{array}\right]$$

Hence the equation becomes

$$2A - 3\left[\begin{array}{rr} 1 & -1 \\ 2 & 1 \end{array}\right] = \left[\begin{array}{rr} 2 & 3 \\ -1 & 2 \end{array}\right]$$

Thus $2A = \left[\begin{array}{rr} 2 & 3 \\ -1 & 2 \end{array}\right] + 3\left[\begin{array}{rr} 1 & -1 \\ 2 & 1 \end{array}\right] = \left[\begin{array}{rr} 5 & 0 \\ 5 & 5 \end{array}\right]$, so finally $A = \frac{1}{2}\left[\begin{array}{rr} 5 & 0 \\ 5 & 5 \end{array}\right] = \frac{5}{2}\left[\begin{array}{rr} 1 & 0 \\ 1 & 1 \end{array}\right]$.

Note that Example 2.1.10 can also be solved by first transposing both sides, then solving for A^T, and so obtaining $A = (A^T)^T$. The reader should do this.

The matrix $D = \left[\begin{array}{rr} 1 & 2 \\ 2 & 5 \end{array}\right]$ in Example 2.1.9 has the property that $D = D^T$. Such matrices are important; a matrix A is called **symmetric** if $A = A^T$. A symmetric matrix A is necessarily square (if A is $m \times n$, then A^T is $n \times m$, so $A = A^T$ forces $n = m$). The name comes from the fact that these matrices exhibit a symmetry about the main diagonal. That is, entries that are directly across the main diagonal from each other are equal.

For example, $\begin{bmatrix} a & b & c \\ b' & d & e \\ c' & e' & f \end{bmatrix}$ is symmetric when $b = b'$, $c = c'$, and $e = e'$.

Example 2.1.11

If A and B are symmetric $n \times n$ matrices, show that $A + B$ is symmetric.

Solution. We have $A^T = A$ and $B^T = B$, so, by Theorem 2.1.2, we have $(A+B)^T = A^T + B^T = A + B$. Hence $A + B$ is symmetric.

Example 2.1.12

Suppose a square matrix A satisfies $A = 2A^T$. Show that necessarily $A = 0$.

Solution. If we iterate the given equation, Theorem 2.1.2 gives

$$A = 2A^T = 2\left[2A^T\right]^T = 2\left[2(A^T)^T\right] = 4A$$

Subtracting A from both sides gives $3A = 0$, so $A = \frac{1}{3}(0) = 0$.

Exercises for 2.1

Exercise 2.1.1 Find a, b, c, and d if

a. $\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c-3d & -d \\ 2a+d & a+b \end{bmatrix}$

b. $\begin{bmatrix} a-b & b-c \\ c-d & d-a \end{bmatrix} = 2\begin{bmatrix} 1 & 1 \\ -3 & 1 \end{bmatrix}$

c. $3\begin{bmatrix} a \\ b \end{bmatrix} + 2\begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

d. $\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} b & c \\ d & a \end{bmatrix}$

Exercise 2.1.2 Compute the following:

a. $\begin{bmatrix} 3 & 2 & 1 \\ 5 & 1 & 0 \end{bmatrix} - 5\begin{bmatrix} 3 & 0 & -2 \\ 1 & -1 & 2 \end{bmatrix}$

b. $3\begin{bmatrix} 3 \\ -1 \end{bmatrix} - 5\begin{bmatrix} 6 \\ 2 \end{bmatrix} + 7\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

c. $\begin{bmatrix} -2 & 1 \\ 3 & 2 \end{bmatrix} - 4\begin{bmatrix} 1 & -2 \\ 0 & -1 \end{bmatrix} + 3\begin{bmatrix} 2 & -3 \\ -1 & -2 \end{bmatrix}$

d. $\begin{bmatrix} 3 & -1 & 2 \end{bmatrix} - 2\begin{bmatrix} 9 & 3 & 4 \end{bmatrix} + \begin{bmatrix} 3 & 11 & -6 \end{bmatrix}$

e. $\begin{bmatrix} 1 & -5 & 4 & 0 \\ 2 & 1 & 0 & 6 \end{bmatrix}^T$

f. $\begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & -4 \\ -2 & 4 & 0 \end{bmatrix}^T$

g. $\begin{bmatrix} 3 & -1 \\ 2 & 1 \end{bmatrix} - 2\begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}^T$

h. $3\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}^T - 2\begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$

Exercise 2.1.3 Let $A = \begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 3 & -1 & 2 \\ 0 & 1 & 4 \end{bmatrix}$, $C = \begin{bmatrix} 3 & -1 \\ 2 & 0 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 3 \\ -1 & 0 \\ 1 & 4 \end{bmatrix}$, and $E = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

Compute the following (where possible).

a. $3A - 2B$
b. $5C$
c. $3E^T$
d. $B + D$
e. $4A^T - 3C$
f. $(A + C)^T$
g. $2B - 3E$
h. $A - D$
i. $(B - 2E)^T$

Exercise 2.1.4 Find A if:

a. $5A - \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} = 3A - \begin{bmatrix} 5 & 2 \\ 6 & 1 \end{bmatrix}$

b. $3A - \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 5A - 2\begin{bmatrix} 3 \\ 0 \end{bmatrix}$

Exercise 2.1.5 Find A in terms of B if:

a. $A + B = 3A + 2B$
b. $2A - B = 5(A + 2B)$

Exercise 2.1.6 If X, Y, A, and B are matrices of the same size, solve the following systems of equations to obtain X and Y in terms of A and B.

a. $5X + 3Y = A$
$2X + Y = B$

b. $4X + 3Y = A$
$5X + 4Y = B$

Exercise 2.1.7 Find all matrices X and Y such that:

a. $3X - 2Y = \begin{bmatrix} 3 & -1 \end{bmatrix}$
b. $2X - 5Y = \begin{bmatrix} 1 & 2 \end{bmatrix}$

Exercise 2.1.8 Simplify the following expressions where A, B, and C are matrices.

a. $2[9(A - B) + 7(2B - A)]$
$-2[3(2B + A) - 2(A + 3B) - 5(A + B)]$

b. $5[3(A - B + 2C) - 2(3C - B) - A]$
$+2[3(3A - B + C) + 2(B - 2A) - 2C]$

Exercise 2.1.9 If A is any 2×2 matrix, show that:

a. $A = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ for some numbers a, b, c, and d.

b. $A = p\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + q\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + r\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + s\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ for some numbers p, q, r, and s.

Exercise 2.1.10 Let $A = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 & 2 \end{bmatrix}$, and $C = \begin{bmatrix} 3 & 0 & 1 \end{bmatrix}$. If $rA + sB + tC = 0$ for some scalars r, s, and t, show that necessarily $r = s = t = 0$.

Exercise 2.1.11

a. If $Q + A = A$ holds for every $m \times n$ matrix A, show that $Q = 0_{mn}$.

b. If A is an $m \times n$ matrix and $A + A' = 0_{mn}$, show that $A' = -A$.

Exercise 2.1.12 If A denotes an $m \times n$ matrix, show that $A = -A$ if and only if $A = 0$.

Exercise 2.1.13 A square matrix is called a **diagonal** matrix if all the entries off the main diagonal are zero. If A and B are diagonal matrices, show that the following matrices are also diagonal.

a. $A + B$
b. $A - B$
c. kA for any number k

Exercise 2.1.14 In each case determine all s and t such that the given matrix is symmetric:

a. $\begin{bmatrix} 1 & s \\ -2 & t \end{bmatrix}$
b. $\begin{bmatrix} s & t \\ st & 1 \end{bmatrix}$
c. $\begin{bmatrix} s & 2s & st \\ t & -1 & s \\ t & s^2 & s \end{bmatrix}$
d. $\begin{bmatrix} 2 & s & t \\ 2s & 0 & s+t \\ 3 & 3 & t \end{bmatrix}$

Exercise 2.1.15 In each case find the matrix A.

a. $\left(A + 3\begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 4 \end{bmatrix}\right)^T = \begin{bmatrix} 2 & 1 \\ 0 & 5 \\ 3 & 8 \end{bmatrix}$

b. $\left(3A^T + 2\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\right)^T = \begin{bmatrix} 8 & 0 \\ 3 & 1 \end{bmatrix}$

c. $\left(2A - 3\begin{bmatrix} 1 & 2 & 0 \end{bmatrix}\right)^T = 3A^T + \begin{bmatrix} 2 & 1 & -1 \end{bmatrix}^T$

d. $\left(2A^T - 5\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}\right)^T = 4A - 9\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$

Exercise 2.1.16 Let A and B be symmetric (of the same size). Show that each of the following is symmetric.

a. $(A - B)$

b. kA for any scalar k

Exercise 2.1.17 Show that $A + A^T$ and AA^T are symmetric for *any* square matrix A.

Exercise 2.1.18 If A is a square matrix and $A = kA^T$ where $k \neq \pm 1$, show that $A = 0$.

Exercise 2.1.19 In each case either show that the statement is true or give an example showing it is false.

a. If $A + B = A + C$, then B and C have the same size.

b. If $A + B = 0$, then $B = 0$.

c. If the $(3,\ 1)$-entry of A is 5, then the $(1,\ 3)$-entry of A^T is -5.

d. A and A^T have the same main diagonal for every matrix A.

e. If B is symmetric and $A^T = 3B$, then $A = 3B$.

f. If A and B are symmetric, then $kA + mB$ is symmetric for any scalars k and m.

Exercise 2.1.20 A square matrix W is called **skew-symmetric** if $W^T = -W$. Let A be any square matrix.

a. Show that $A - A^T$ is skew-symmetric.

b. Find a symmetric matrix S and a skew-symmetric matrix W such that $A = S + W$.

c. Show that S and W in part (b) are uniquely determined by A.

Exercise 2.1.21 If W is skew-symmetric (Exercise 2.1.20), show that the entries on the main diagonal are zero.

Exercise 2.1.22 Prove the following parts of Theorem 2.1.1.

a. $(k + p)A = kA + pA$

b. $(kp)A = k(pA)$

Exercise 2.1.23 Let A, A_1, A_2, $\dots$, A_n denote matrices of the same size. Use induction on n to verify the following extensions of properties 5 and 6 of Theorem 2.1.1.

a. $k(A_1 + A_2 + \cdots + A_n) = kA_1 + kA_2 + \cdots + kA_n$ for any number k

b. $(k_1 + k_2 + \cdots + k_n)A = k_1A + k_2A + \cdots + k_nA$ for any numbers k_1, k_2, $\dots$, k_n

Exercise 2.1.24 Let A be a square matrix. If $A = pB^T$ and $B = qA^T$ for some matrix B and numbers p and q, show that either $A = 0 = B$ or $pq = 1$.
[*Hint*: Example 2.1.7.]

2.2 Matrix-Vector Multiplication

Up to now we have used matrices to solve systems of linear equations by manipulating the rows of the augmented matrix. In this section we introduce a different way of describing linear systems that makes more use of the coefficient matrix of the system and leads to a useful way of "multiplying" matrices.

Vectors

It is a well-known fact in analytic geometry that two points in the plane with coordinates $(a_1,\ a_2)$ and $(b_1,\ b_2)$ are equal if and only if $a_1 = b_1$ and $a_2 = b_2$. Moreover, a similar condition applies to points $(a_1,\ a_2,\ a_3)$ in space. We extend this idea as follows.

An ordered sequence $(a_1,\ a_2,\ \dots,\ a_n)$ of real numbers is called an **ordered *n*-tuple**. The word "ordered" here reflects our insistence that two ordered n-tuples are equal if and only if corresponding entries are the same. In other words,

$$(a_1,\ a_2,\ \dots,\ a_n) = (b_1,\ b_2,\ \dots,\ b_n) \quad \text{if and only if} \quad a_1 = b_1,\ a_2 = b_2,\ \dots,\ \text{ and } a_n = b_n.$$

Thus the ordered 2-tuples and 3-tuples are just the ordered pairs and triples familiar from geometry.

Definition 2.4 The set $\mathbb{R}^n$ of ordered n-tuples of real numbers

Let $\mathbb{R}$ denote the set of all real numbers. The set of all *ordered n-tuples from $\mathbb{R}$ has a special notation:*

$\mathbb{R}^n$ denotes the set of all ordered n-tuples of real numbers.

There are two commonly used ways to denote the n-tuples in $\mathbb{R}^n$: As rows $(r_1,\ r_2,\ \dots,\ r_n)$ or columns $\left[\begin{array}{c} r_1 \\ r_2 \\ \vdots \\ r_n \end{array} \right]$; the notation we use depends on the context. In any event they are called **vectors** or ***n*-vectors** and will be denoted using bold type such as $\mathbf{x}$ or $\mathbf{v}$. For example, an $m \times n$ matrix A will be written as a row of columns:

$$A = \left[\begin{array}{cccc} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{array} \right] \text{ where } \mathbf{a}_j \text{ denotes column } j \text{ of } A \text{ for each } j.$$

If $\mathbf{x}$ and $\mathbf{y}$ are two n-vectors in $\mathbb{R}^n$, it is clear that their matrix sum $\mathbf{x} + \mathbf{y}$ is also in $\mathbb{R}^n$ as is the scalar multiple $k\mathbf{x}$ for any real number k. We express this observation by saying that $\mathbb{R}^n$ is **closed** under addition and scalar multiplication. In particular, all the basic properties in Theorem 2.1.1 are true of these n-vectors. These properties are fundamental and will be used frequently below without comment. As for matrices in general, the $n \times 1$ zero matrix is called the **zero *n*-vector** in $\mathbb{R}^n$ and, if $\mathbf{x}$ is an n-vector, the n-vector $-\mathbf{x}$ is called the **negative x**.

Of course, we have already encountered these n-vectors in Section 1.3 as the solutions to systems of linear equations with n variables. In particular we defined the notion of a linear combination of vectors and showed that a linear combination of solutions to a homogeneous system is again a solution. Clearly, a linear combination of n-vectors in $\mathbb{R}^n$ is again in $\mathbb{R}^n$, a fact that we will be using.