Definition 2.6: Scalar-vector multiplication

Let $\mathbf{x} \in \mathbb{R}^n$ be a column vector and let $a \in \mathbb{R}$ be a scalar. We define the **scalar-vector multiplication** as the vector

$$a\,\mathbf{x} = a \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{bmatrix}$$

where the ith coefficient of this product is given by ax_i for all $i \in \{1, 2, ..., n\}$.

The scalar-vector multiplication $a\mathbf{x} \in \mathbb{R}^n$ is a column vector. The left argument of the product is a scalar a and the right argument is a vector $\mathbf{x}$. We see that scalar-vector multiplication is a function from $\mathbb{R} \times \mathbb{R}^n$ to $\mathbb{R}^n$.

EXAMPLE 2.2.1

Let's return to our mass-spring system from Example 2.1.5 in which we measured the relationship between the mass placed on the movable end of a spring and the position of that end. The data vector $\mathbf{m}_{\text{obs}}$ stored the masses hung on the free end of our spring. **Newton's Second Law** states that the force acting on an object is calculated as the mass of that object (measured in kilograms) times the acceleration of that object (measured in meters per second squared). Mathematically, we write the scalar equation

$$f = m\,a = m\,\ddot{x}(t)$$

where $\ddot{x}(t) = \dfrac{d^2}{dt^2}[x(t)]$ and $x(t)$ is the position function for the object to which we apply our force.

To find the force vector $\mathbf{f}_{\text{calc}} \in \mathbb{R}^{3\times 1}$ corresponding to the mass vector $\mathbf{m}_{\text{obs}}$, we use scalar-vector multiplication. The ith entry of our force vector is given by $f_i = am_i$, where $a = 9.8\text{m/s}^2$ is the constant of acceleration due to earth's gravity (we assumed that we conducted this experiment on the face of the earth). Our calculated force vector corresponding to the masses used in this experiment is given by the scalar-vector product:

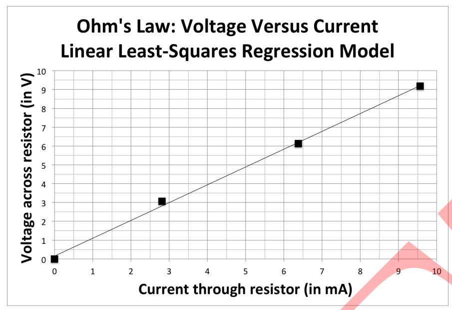
$$\mathbf{f}_{\text{calc}} = 9.8\,\mathbf{m}_{\text{obs}} = 9.8 \begin{bmatrix} 0.00000 \\ 0.20010 \\ 0.40031 \end{bmatrix} = \begin{bmatrix} 0.00000 \\ 1.96098 \\ 3.92304 \end{bmatrix}$$

EXAMPLE 2.2.2

Let's study the mathematical relationship between the current running through a resistor and the voltage across that resistor. Look back on Example 2.1.6 where we collected the data vectors

$$\mathbf{v}_{\text{obs}} = \begin{bmatrix} 0.00 \\ 3.06 \\ 6.13 \\ 9.18 \end{bmatrix}, \qquad \mathbf{i}_{\text{obs}} = \begin{bmatrix} 0.00 \\ 2.81 \\ 6.38 \\ 9.57 \end{bmatrix}.$$

The experimental data seems to demonstrate a linear relationship. We superimpose a "line of best fit" on the graph of our data points to test our intuition:



Although we see clear linear relationship, we notice that not all data points lie exactly on this line. The discrepancies between our model and our observed data are due to experimental errors. We discuss such errors in our development of techniques to solve least-squares problems. In fact, the equation for the line of best fit that you see in this diagram comes from the solution to a least-squares problem. For now, we use a multimeter to measure the resistance of our resistor in our experiment and find $r = 0.990k\Omega = 990\Omega$. We substitute this value for r into the vector version of **Ohm's Law**, given by

$$\mathbf{v}_{\text{obs}} = r\,\mathbf{i}_{\text{calc}} \tag{2.2}$$

where $\mathbf{v}_{\text{obs}} \in \mathbb{R}^4$ is the vector of measured voltages across our resistor and $r \in \mathbb{R}$ is measured resistance for the resistor in our experiment. We calculate the vector $\mathbf{i}_{\text{calc}} \in \mathbb{R}^4$ using scalar-vector multiplication to find

$$\mathbf{i}_{\text{calc}} = \frac{1}{.990} \begin{bmatrix} 0.00 \\ 3.06 \\ 6.13 \\ 9.18 \end{bmatrix} = \begin{bmatrix} 0.00 \\ 3.09 \\ 6.19 \\ 9.27 \end{bmatrix}.$$

We expect approximations to be slightly different from our observed currents stored in $\mathbf{i}_{\text{obs}}$ due to experimental error.

We can interpret scalar-vector multiplication geometrically to represents the "stretching" of a vector.

EXAMPLE 2.2.3

Consider the vector $\mathbf{x} \in \mathbb{R}^2$ define by

$$\mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

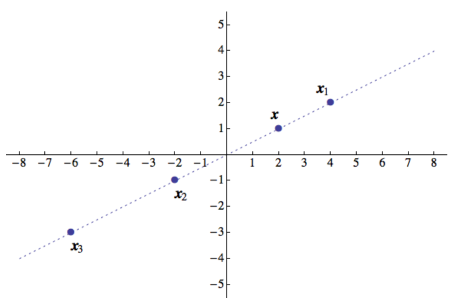
Suppose we are using a vector model in which individual vectors represents single points in $\mathbb{R}^2$. We make this vector longer or shorter by multiplying by the appropriate scalar using the scalar-vector multiplication

$$\alpha \mathbf{x}$$

For example, suppose that

$$\mathbf{x}_1 = 2 \cdot \mathbf{x} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, \qquad \mathbf{x}_2 = -1 \cdot \mathbf{x} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}, \qquad \mathbf{x}_3 = -3 \cdot \mathbf{x} = \begin{bmatrix} -6 \\ -3 \end{bmatrix},$$

We can graph each of these vectors on the same axis.



We see that changing the value of the scalar α scales the first and second coordinates of the vector $\mathbf{x}$ by the same value. This effectively shifts the original point along the line $2y = x$. As we will see, we call the line $2y = x$ the span of the vector $\mathbf{x}$.

Definition 2.7: Column Vector Addition

Given two column vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we define an operation on these two vectors known as column vector addition as

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

When summing two vectors $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$, we must have that $n = m$. In other words, we can only sum vectors that have the same number of rows and the same number of columns. The ith coefficient of the sum of the vectors $\mathbf{x}$ and $\mathbf{y}$ is simply the sum of the ith coefficients of $\mathbf{x}$ and $\mathbf{y}$, for $i = 1, 2, ..., n$.

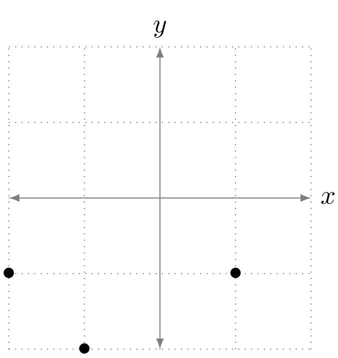
Notice that the sum of two column vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ is also a column vector $\mathbf{x} + \mathbf{y} \in \mathbb{R}^n$. Thus, we can conclude that column vector addition is a relation from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}^n$. Column vector addition can be used to combine information from two column vectors to form a new column vector.

EXAMPLE 2.2.4

Recall our triangle from example 2.1.1. Recall that the three vertices of our triangle were given by We can use column vectors to define the vertices of a triangle in the plane. For example, consider the three column vectors

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \qquad \mathbf{v}_3 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Suppose we want to create a second triangle with three new vertices by shifting the vertices $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ to the left by one unit and down by two units. We can accomplish this using vector addition. In particular, let



$$\mathbf{s} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$$

Then, if we define $\mathbf{x}_i = \mathbf{v}_i + \mathbf{s}$ we see

$$\mathbf{x}_1 = \begin{bmatrix} -1 \\ -2 \end{bmatrix}, \qquad \mathbf{x}_2 = \begin{bmatrix} -2 \\ -1 \end{bmatrix}, \qquad \mathbf{x}_3 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Notice that the vertices of our new triangle are indeed shifted left one and down two compared with our original vertices.

EXAMPLE 2.2.5

Look back at Example 2.1.5. The vector $\mathbf{x} \in \mathbb{R}^{3\times 1}$ encoded the position of the movable end of the spring as measured using our experiment apparatus. However, when studying the relationship between the internal force of the spring, we want to focus on the displacement of the moveable end. That is, we want calculate the difference between the position of the movable end under our various forces and the initial position under no force. We can use scalar-vector multiplication to encode this data as a displacement vector:

$$\mathbf{u} = \mathbf{x}_{\text{obs}} - \mathbf{x}_0 = \begin{bmatrix} 1.040 \\ 0.932 \\ 0.820 \end{bmatrix} - \begin{bmatrix} 1.040 \\ 1.040 \\ 1.040 \end{bmatrix} = \begin{bmatrix} 0.000 \\ -0.108 \\ -0.220 \end{bmatrix}$$

In this equation, we have $\mathbf{x}_0 \in \mathbb{R}^{3\times 1}$ is the vector containing entries equal to the initial position 1.040m. The ith coefficient of vector $\mathbf{u}$ gives the difference between the initial position of the free end of the spring with no mass and end the end position of the spring when mass m_i is attached.

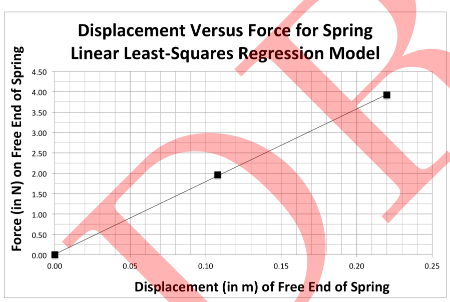
EXAMPLE 2.2.6

In studying the physical properties of mass-spring systems, we often use **Hooke's law** which states that the internal forces of a spring are directly proportional to the elongation of the spring. We can state this in vector form as follows:

$$\mathbf{f} = -k\mathbf{u}$$

where k is the specific spring constant for the spring we use in our experiment, $\mathbf{u}$ is the calculated displacement vector from Example 2.2.5 and $\mathbf{f}$ is the negative of the calculated force vector from Example 2.2.1.

Note that the spring constant of a spring is a measurement of stiffness. The higher the spring constant, the harder it is to pull the spring apart. We can use excel's trend line chart option to get a formula for the value of k in this experiment (see the figure below):



There are a number of algebraic properties that hold for these operations on column vectors.

Theorem 5: Algebraic Properties of Vector Addition and Scalar Multiplication

Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^{n\times 1}$ and $a, b \in \mathbb{R}$. Then, all of the following are properties of column vector addition:

1. Commutativity of vector addition: $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}$
2. Associativity of vector addition: $\mathbf{x}+(\mathbf{y}+\mathbf{z})=(\mathbf{x}+\mathbf{y})+\mathbf{z}$
3. Additive identity: $\mathbf{x}+\mathbf{0}=\mathbf{0}+\mathbf{x}=\mathbf{x}$
4. Additive inverses: $\mathbf{x}+-\mathbf{x}=-\mathbf{x}+\mathbf{x}=\mathbf{0}$ with $-\mathbf{x}=-1\mathbf{x}$
5. Distributivity over vector addition: $a(\mathbf{x}+\mathbf{y})=a\mathbf{x}+a\mathbf{y}$
6. Distributivity over scalar addition: $(a+b)\mathbf{x}=a\mathbf{x}+b\mathbf{x}$
7. Associativity of scalar multiplication: $a\,(b\mathbf{x})=(ab)\mathbf{x}$
8. Multiplicative identity of scalar multiplication: $1\mathbf{x}=\mathbf{x}$

Prove each one of these using a step-by-step rigorous format.

Proof. To prove each of these statements, we begin by letting $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^{n\times 1}$ and $a, b \in \mathbb{R}$. We will then work through each property, one-by-one, to reach our desired conclusions. Let's start with commutativity. Consider

$$\mathbf{x}+\mathbf{y}=\begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix}$$

However, by definition we know that $x_i, y_i \in \mathbb{R}$ for all $i \in \{1, 2, ..., n\}$ and thus we know that $x_i+y_i=y_i+x_i$ since addition of real numbers is commutative. Since this holds true for all index values i, we see that

$$\mathbf{x}+\mathbf{y}=\begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix}=\begin{bmatrix} y_1+x_1 \\ y_2+x_2 \\ \vdots \\ y_n+x_n \end{bmatrix}=\mathbf{y}+\mathbf{x}.$$

Next, let's prove associativity. Consider

$$\mathbf{x}+(\mathbf{y}+\mathbf{z})=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}+\begin{bmatrix} y_1+z_1 \\ y_2+z_2 \\ \vdots \\ y_n+z_n \end{bmatrix}=\begin{bmatrix} x_1+(y_1+z_1) \\ x_2+(y_2+z_2) \\ \vdots \\ x_n+(y_n+z_n) \end{bmatrix}$$

Again, we can focus on the scalar addition $x_i+(y_i+z_i)$ for any choice $i \in \{1, 2, ..., n\}$. We know that addition of real numbers is associative and thus we have

$$x_i + (y_i + z_i) = (x_i + y_i) + z_i.$$

Since this holds for all index values we see

$$\mathbf{x} + (\mathbf{y} + \mathbf{z}) = \begin{bmatrix} x_1 + (y_1 + z_1) \\ x_2 + (y_2 + z_2) \\ \vdots \\ x_n + (y_n + z_n) \end{bmatrix} = \begin{bmatrix} (x_1 + y_1) + z_1 \\ (x_2 + y_2) + z_2 \\ \vdots \\ (x_n + y_n) + z_n \end{bmatrix} = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$$

and we have confirmed that vector addition is associative.

We continue by proving our desired property for the additive identity $\mathbf{0} \in \mathbb{R}^n$. Consider

$$\mathbf{x} + \mathbf{0} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 + 0 \\ x_2 + 0 \\ \vdots \\ x_n + 0 \end{bmatrix}$$

Since each coefficient $x_i + 0 = x_i$ we conclude that $\mathbf{x} + \mathbf{0} = \mathbf{x}$. By similar reasoning we can conclude that $\mathbf{x} = \mathbf{0} + \mathbf{x}$ which completes our proof. $\square$

Each of the rest of the Algebraic Properties of Vector Addition and Scalar Multiplication can be proven using the same types of arguments outlined above. The reader is encouraged to complete these proofs as part of your exercise set associated with this section.

One of the major take away points from these algebraic properties is that vectors in $\mathbb{R}^n$ borrow much of their algebraic structure from the real numbers. As we see in the proofs above, this is a direct consequence of the fact that each coefficients lives in the real numbers and thus individual operations on each coefficient satisfy all algebraic properties of real numbers. One of the most fundamental and powerful attributes of vectors is the ability to represent numerous operations in a very compact form. For given vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we see that the vector-vector addition $\mathbf{x}+\mathbf{y}$ represents n different additions being executed simultaneously. Similarly, For vector $\mathbf{x} \in \mathbb{R}^n$ and scalar $a \in \mathbb{R}$, the scalar-vector multiplication $a\mathbf{x}$ represents a total of n scalar-scalar multiplications. With this in mind, think back to arithmetic.

If we were required to write down each operation one-by-one for vectors of size $n = 100$, this means we would have to track 100 different operations and carry out each operation individually. However, vector notation and vector operations allow us to suppress the details of these calculations and instead focus on producing relevant new vectors. This is a subtle yet beautiful property of linear algebra: we define multidimensional data using individual symbols, create relevant operations to manipulate this data and use very slick notation to hide the arithmetic details and focus on the larger structures to accomplish our goals.

Definition 2.8: The Transpose of a vector

Let $\mathbf{x} \in \mathbb{R}^n$ be a column vector. We define the **transpose** of $\mathbf{x}$ as the $1 \times n$ row vector

$$\mathbf{x}^T = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$$

where the entry in the ith row of $\mathbf{x}$ becomes the entry in the ith column of $\mathbf{x}^T$.

On the other hand, let $\mathbf{y} \in \mathbb{R}^{1\times n}$ be a row vector. The **transpose** of $\mathbf{y}$ is the $n \times 1$ column vector

$$\mathbf{y}^T = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

where the coefficient in the kth column of $\mathbf{y}$ becomes the entry in the kth row of $\mathbf{y}^T$.

When taking the transpose of a vector, we switch the row and column indices. This implies that the transpose of row vectors are column vectors while the transpose of column vectors are row vectors.

Theorem 6: Algebraic Properties of Vector Transposes

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{n\times 1}$ and $a, b \in \mathbb{R}$. Then, all of the following are properties of column vector addition:

i. $\left(\mathbf{x}^T\right)^T = \mathbf{x}$

ii. $\left(\mathbf{x} + \mathbf{y}\right)^T = \mathbf{x}^T + \mathbf{y}^T$

iii. $\left(a\,\mathbf{x}\right)^T = a\,\mathbf{x}^T$

Proof. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $a \in \mathbb{R}$. We begin by proving our first property. Consider.

$$\left(\mathbf{x}^T\right)^T = \left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}^T\right)^T = \left(\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}\right)^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{x}.$$

This is exactly what we wanted to show.

We continued by proving that the transpose of the sum of vectors is the sum of the transposes of each vector. Consider

$$
\begin{aligned}
(\mathbf{x}+\mathbf{y})^T &= \left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \right)^T \\
&= \left(\begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix} \right)^T \\
&= \begin{bmatrix} x_1 + y_1 & x_2 + y_2 & \cdots & x_n + y_n \end{bmatrix} \\
&= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} + \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} \\
&= \mathbf{x}^T + \mathbf{y}^T.
\end{aligned}
$$

This is what we wanted to show.

Finally, let's establish that scalar multiplication goes through the transpose operation:

$$
\begin{aligned}
(a\mathbf{x})^T &= \left(a \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right)^T \\
&= \left(\begin{bmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{bmatrix} \right)^T \\
&= \begin{bmatrix} ax_1 & ax_2 & \cdots & ax_n \end{bmatrix} \\
&= a \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \\
&= a\mathbf{x}^T.
\end{aligned}
$$

Here we are done with our proof. □

Although the above theorem and proof are stated in terms of column vectors, the same properties hold if we assume that $\mathbf{x}, \mathbf{y}$ are row vectors. As we will see later, the transpose operation is extremely helpful in re-interpreting dot products between vectors as matrix-matrix multiplication. Transposes are also helpful in formulating outer-products between vectors.

Lesson 4 Exercises: Vectors Operations

1. Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^4$ be given by

$$\mathbf{x} = \begin{bmatrix} 2 \\ 3 \\ -1 \\ 6 \end{bmatrix}, \qquad \mathbf{y} = \begin{bmatrix} 4 \\ 0 \\ -2 \\ -1 \end{bmatrix}, \qquad \mathbf{z} = \begin{bmatrix} -1 \\ 1 \\ 0 \\ -5 \end{bmatrix},$$

   Use these vectors to find each of the following:

   a. $\mathbf{x} + \mathbf{z}$
   b. $\mathbf{x}^T + \mathbf{y}^T$
   c. $\mathbf{y} + \mathbf{z}^T$
   d. $2\mathbf{x} - 1\mathbf{y}$

2. Prove Theorem 6: The Algebraic Properties of Vector Addition and Scalar Multiplication for vectors in $\mathbb{R}^n$

3. Prove Theorem 7: The Algebraic Properties of Vector Transposes in $\mathbb{R}^n$

4. Create a vector model for some data that you can collect in your daily life. Discuss how to use scalar-vector multiplication and vector addition to analyze this data.

5. Analyze all possible geometric transformations of the vertices of a triangle that can be accomplished using only scalar-vector multiplication and vector addition.