Chapter 1

$\mathbb{R}^n$

A matrix is simply an array of numbers, but they have certain algebraic properties which we will discuss. An important interpretation of matrices is that of functions on $\mathbb{R}^n$, and many of the properties make the most sense with that interpretation. So we will begin discussion $\mathbb{R}^n$ before we get to matrices. You may recall $\mathbb{R}^n$ from Calculus, but a brief reminder may be helpful.

Definition. The set $\mathbb{R}^n$ is the set of all n-tuples of real numbers, written as a column. $\square$

So, for example, $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is an element of $\mathbb{R}^2$, and $\begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix}$ is an element of $\mathbb{R}^3$. When you learned about $\mathbb{R}^n$ in Calculus, the elements were probably given as rows rather than columns; an element of $\mathbb{R}^2$ would be written (x, y) rather than $\begin{pmatrix} x \\ y \end{pmatrix}$. While you should be familiar with both ways of representing elements of $\mathbb{R}^n$, when dealing with matrices it is usually more useful to think of elements of $\mathbb{R}^n$ as columns.

An element of $\mathbb{R}^n$ is often represented by $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, although, to avoid subscripts, elements of $\mathbb{R}^2$ are usually represented as $\begin{pmatrix} x \\ y \end{pmatrix}$, and elements of $\mathbb{R}^3$ by $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$. The numbers x_i are called **coordinates**.

Recall that an element of $\mathbb{R}^2$ can be associated with a point in the plane (see figure 1.1). For our purposes, it will be useful to think of

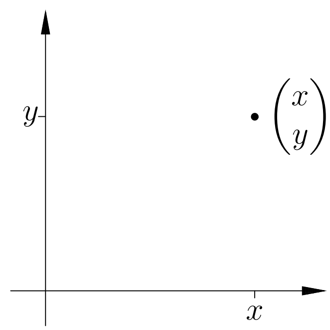

Figure 1.1: *An element of* $\mathbb{R}^2$ *is associated with a point in the plane.*

an element of $\mathbb{R}^2$ as directions from the origin to the point; the direction and distance can be represented by an arrow from the origin to the point (see figure 1.2). Similarly, and element of $\mathbb{R}^3$ can be represented

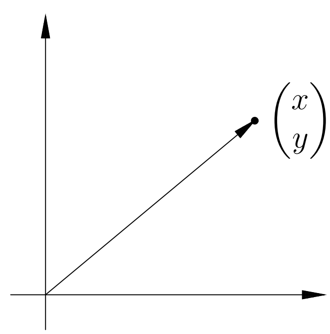

Figure 1.2: *An element of* $\mathbb{R}^2$ *can be associated with an arrow.*

by a point in space, or as an arrow indicating a direction and distance in space. Elements of $\mathbb{R}^n$ for $n > 3$ can also be associated with points and arrows in higher dimensional spaces, although they get harder to visualize. Viewed this way, elements of $\mathbb{R}^n$ are called ***vectors***.

Vectors are often described as having **length** and **direction**. The direction can be described by simply giving the vector itself, although there other ways to do it. The length of a vector is a number. The Pythagorean theorem leads to the following definition.

Definition. The ***length*** of a vector $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ is

$$\|\boldsymbol{x}\| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

□

Example. The length of $\boldsymbol{x} = \begin{pmatrix} 3 \\ 4 \\ 4 \end{pmatrix}$ is

$$\|\boldsymbol{x}\| = \sqrt{3^2 + 4^2 + 5^2} = \sqrt{50} = 5\sqrt{2}.$$

☑

If you've had multivariable Calculus, you may recall that elements of $\mathbb{R}^n$ can be added. Two vectors can be added by adding their corresponding coordinates.

Example. The sum of the elements $\begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}$ and $\begin{pmatrix} 3 \\ 1 \\ 7 \end{pmatrix}$ of $\mathbb{R}^3$ is

$$\begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \\ 7 \end{pmatrix} = \begin{pmatrix} 2+3 \\ 3+1 \\ 5+7 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \\ 12 \end{pmatrix}.$$

☑

The sum $\boldsymbol{x} + \boldsymbol{y}$ of two vectors in $\mathbb{R}^n$ can be thought of as following the directions given by $\boldsymbol{x}$ followed by following the directions given by $\boldsymbol{y}$. Geometrically, to add $\boldsymbol{x} + \boldsymbol{y}$, take the tail of an arrow representing $\boldsymbol{y}$ and place it on the tip of an arrow representing $\boldsymbol{x}$ (see figure 1.3). The arrow representing $\boldsymbol{x} + \boldsymbol{y}$ is then the arrow from the tail of the $\boldsymbol{x}$ arrow to the tip of the $\boldsymbol{y}$ arrow (see figure 1.4).

Notice that $\mathbb{R}^n$ has the special element $\mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ which is the additive identity: $v + \mathbf{0} = v$ for any $v \in \mathbb{R}^n$.

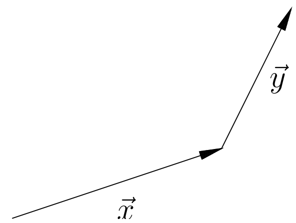

Figure 1.3: *Placing vectors in $\mathbb{R}^n$ tip to tail.*

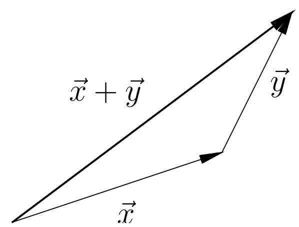

Figure 1.4: *Adding two vectors in $\mathbb{R}^n$.*

Vectors can also be multiplied by numbers by multiplying each coordinate by the number.

Example. The product of 5 and the element $\begin{pmatrix} 2 \\ 1 \\ 2 \\ 3 \end{pmatrix}$ of $\mathbb{R}^4$ is

$$\begin{pmatrix} 5 \cdot 2 \\ 5 \cdot 1 \\ 5 \cdot 2 \\ 5 \cdot 3 \end{pmatrix} = \begin{pmatrix} 10 \\ 5 \\ 10 \\ 15 \end{pmatrix}.$$

☑

Multiplying a vector $\boldsymbol{x}$ by a number c can be thought of as following the directions given by $\boldsymbol{x}$, but going c times as far. Geometrically, multiplying $\boldsymbol{x}$ by c results in an arrow with the same direction as the arrow representing $\boldsymbol{x}$ (or reversing the direction, if c is negative), but changing the length by a factor of c (see figure 1.5).

Proposition. *If $\boldsymbol{x} \in \mathbb{R}^n$ and $c \in \mathbb{R}$, then $\|c\boldsymbol{x}\| = |c|\|\boldsymbol{x}\|$.* □

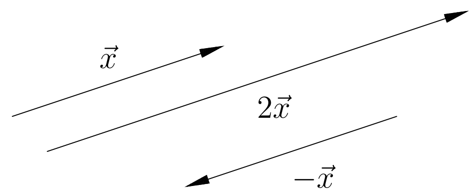

Figure 1.5: *Multiplying a vector by a number.*

Since multiplying a vector by a number serves to scale the length of the vector, this type of multiplication is called ***scalar multiplication***, and numbers are called ***scalars***.

Addition and scalar multiplication are the two ***vector space operations***. Although we will (mostly) restrict our attention to $\mathbb{R}^n$, any set in which there is defined addition and scalar multiplication is called a ***vector space***.

Definition. A ***vector space*** is a set V on which there are two operations:
1. Addition: if $\boldsymbol{x}, \boldsymbol{y} \in V$, then $\boldsymbol{x} + \boldsymbol{y} \in V$.
2. Scalar multiplication: if $\boldsymbol{x} \in V$ and $c \in \mathbb{R}$, then $c\boldsymbol{x} \in V$.

Addition and scalar multiplication need to satisfy certain naturality conditions, the specifics of which we'll omit. □

An example of a vector space (besides $\mathbb{R}^n$) is the set of all polynomials. Any two polynomials can be added to get another polynomial, and a polynomial can be multiplied by a number to get another polynomial. Two polynomials can be multiplied to get another polynomial, but that is irrelevant when we are considering polynomials as a vector space.

Even when we restrict our attention to $\mathbb{R}^n$, we will come across more vector spaces than $\mathbb{R}^n$ itself. A subset of $\mathbb{R}^n$ will have a special status if addition and scalar multiplication will preserve it. Specifically:

Definition. A subset S of $\mathbb{R}^n$ is a ***subspace*** if:
1. for any $\vec{x}, \vec{y} \in S$, their sum $\vec{x} + \vec{y}$ is also in S. (We say that S is ***closed under addition***.)
2. for any $\vec{x} \in S$, $c \in \mathbb{R}$, the scalar product $c\vec{x}$ is also in S. (We say that S is ***closed under scalar multiplication***.) □

Every $\mathbb{R}^n$ has two somewhat trivial subspaces. The entire space $\mathbb{R}^n$ is a subspace of itself, and the one element subset $\{\mathbf{0}\}$ is also a subspace of $\mathbb{R}^n$. Every other subspace is between these two: If S is a subspace of $\mathbb{R}^n$, then by definition S must be contained in $\mathbb{R}^n$, but also the element $\vec{0}$ must be an element of S.

Example. The set of vectors in $\mathbb{R}^3$ whose second component is 0 (i.e., elements of the form $\begin{pmatrix} x \\ 0 \\ z \end{pmatrix}$) is a subspace of $\mathbb{R}^3$. This is straightforward to check:

1. We can add two elements of this type
$$\begin{pmatrix} x_1 \\ 0 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ 0 \\ z_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ 0 \\ z_1 + z_2 \end{pmatrix}$$
to get another element of this type. So this set is closed under addition.
2. We can multiply an element of this type by a number
$$c \begin{pmatrix} x \\ 0 \\ z \end{pmatrix} = \begin{pmatrix} cx \\ 0 \\ cz \end{pmatrix}$$
to get another element of this type. So this set is closed under scalar multiplication.

So this is indeed a subspace. ☑

Example. The set of vectors in $\mathbb{R}^4$ whose first element is 1 (i.e., elements of the form $\begin{pmatrix} 1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$ is not a subspace of $\mathbb{R}^4$, since the zero element $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ is not in this set. ☑

Example. The set of vectors in $\mathbb{R}^3$ in which the first or second component is 0 (i.e., the set of vectors of the form $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ where either $x = 0$ or $z = 0$) is not a subspace of $\mathbb{R}^3$. To see this, note that we can add two elements of this set and get an element not in this set:

$$\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

so the set is not closed under addition. ☑

The subspaces of $\mathbb{R}^n$ can be characterized as the solution sets of certain systems of equations.

Definition. A ***linear equation*** in n variables is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b.$$ □

Example.

$$2x_1 + x_2 - 4x_3 + x_4 = 5.$$ ☑

Example.

$$x + 2y - z = 7.$$ ☑

Example. The equation

$$y = x - z + 2$$

can be rewritten

$$-x + y + z = 2,$$

and so is linear. ☑

We will be interested in several simultaneous equations. A ***linear system of m equations in n unknowns*** is a system of the form

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = b_2 \\ \vdots & & & & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = b_m \end{array}$$

where all equations are expected to hold.

Example. The system

$$\begin{array}{rcrcrl} x & - & 3y & + & 2z & = 5 \\ x & + & y & + & z & = 1 \\ -3x & & & + & z & = 0 \\ & & y & - & 2x & = 5 \end{array}$$

is a system of 4 equations in 3 unknowns. ☑

Linear equations are an important type of equation, and solving linear equations is an important use of matrices, which we will get to in later chapters.

Definition. A system of linear equations is ***homogeneous*** if the constants are all 0; i.e., it is of the form

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = 0 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = 0 \\ \vdots & & & & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = 0 \end{array}$$ □

Note that the solution set of a homogeneous linear equation is a subspace of the appropriate $\mathbb{R}^n$. Let's show that in a simple case.

Example. The set S of vectors $\begin{pmatrix} x \\ y \end{pmatrix}$ in $\mathbb{R}^2$ with $2x - y = 0$ is a subspace of $\mathbb{R}^2$.

To see that:

1. Suppose $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ and $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ are in S. Then $2x_1 - y_1 = 0$ and $2x_2 - y_2 = 0$. The sum of the vectors, $\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix}$, is also in S, since

$$\begin{aligned} 2(x_1 + x_2) - (y_1 + y_2) &= 2x_1 + 2x_2 - y_1 - y_2 \\ &= (2x_1 - y_1) + (2x_2 - y_2) \\ &= 0 + 0 = 0. \end{aligned}$$

   So S is closed under addition.
2. Suppose $\begin{pmatrix} x \\ y \end{pmatrix}$ is in S and c is a number. Then $2x - y = 0$. The product $\begin{pmatrix} cx \\ cy \end{pmatrix}$ is in S, since

$$2(cx) - (cy) = c(2x - y) = c \cdot 0 = 0.$$

   So S is closed under scalar multiplication.

So S is a subspace. ☑

In fact, we have the following.

Proposition. *A subset of $\mathbb{R}^n$ is a subspace exactly when it is the solution set of a system of homogeneous linear equations.* $\square$

Example. The set of vectors $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ in $\mathbb{R}^3$ with $x = 0$ or $y + 2z = 0$ is the solution set of the homogeneous system

$$\begin{aligned} x \quad\quad\quad &= 0 \\ y \;\; +2x &= 0 \end{aligned}$$

and so it is a subspace of $\mathbb{R}^3$. ☑

If we want to visualize the subspaces of $\mathbb{R}^n$, let's start with $\mathbb{R}^2$.

1. The trivial subspace $\{\mathbf{0}\}$ is itself a subspace.
2. If a subspace contains $\boldsymbol{x} \neq 0$, then it must contain all multiples of $\boldsymbol{x}$, $\{c\boldsymbol{x} : c \in \mathbb{R}\}$. The set $\boldsymbol{x}$, $\{c\boldsymbol{x} : c \in \mathbb{R}\}$ is itself a subspace of $\mathbb{R}^2$, corresponding to a line through the origin.

3. If a subspace contains $\boldsymbol{x} \neq 0$ and $\boldsymbol{y}$ which is not a multiple of $\boldsymbol{x}$, then it will have to be all of $\mathbb{R}^2$.

So, geometrically, the subspaces of $\mathbb{R}^2$ are:

1. the trivial subspace $\{\vec{0}\} = \{\begin{pmatrix} 0 \\ 0 \end{pmatrix}\}$
2. lines through the origin
3. $\mathbb{R}^2$ itself.

Similarly, the subspaces of $\mathbb{R}^3$ are:

1. the trivial subspace $\{\vec{0}\} = \{\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\}$
2. lines through the origin
3. planes through the origin
4. $\mathbb{R}^3$ itself.

One operation on $\mathbb{R}^n$ that you may recall from Calculus is the dot product. We won't be using this for a while, but a brief reminder now couldn't hurt.

Definition. Let $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$. Then the ***dot product*** of $\boldsymbol{x}$ and $\boldsymbol{y}$ is

$$\boldsymbol{x} \cdot \boldsymbol{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

□

Example. Let $\boldsymbol{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $vecty = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}$. Then

$$\boldsymbol{x} \cdot \boldsymbol{y} = 1 \times 2 + 2 \times 1 + 3 \times 4 = 16.$$

☑

Since the dot product of two vectors is not a vector, but rather a number (or scalar), it is also called the scalar product.

The dot product is often written as $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$. Since this notation no longer uses a dot, it is also called the ***inner product***. The inner product has the following properties.

Proposition. *For any $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z} \in \mathbb{R}^n$, and any number a:*

- $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$.
- $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$.
- $\langle a\boldsymbol{x}, \boldsymbol{y} \rangle = a \langle \boldsymbol{x}, \boldsymbol{y} \rangle$.
- $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$, *and* $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ *only if* $\boldsymbol{x} = \mathbf{0}$. □

Remark. Any function $\langle \cdot, \cdot \rangle$ of two vectors which satisfies the properties of the previous proposition is called an inner product. We will stick to using the standard inner product, the one defined above, even though much of what follows would also be true if we used a different inner product. ⊠

Note that if $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$, then $\|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$. So in terms of the inner product, the ***length*** of a vector $\boldsymbol{x}$ is given as

$$\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}.$$

To avoid square roots, we can write $\|\boldsymbol{x}\|^2 = \langle \boldsymbol{x}, \boldsymbol{x} \rangle$.

You may recall that given two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^n$, if the angle between them is θ, then

$$\boldsymbol{x} \cdot \boldsymbol{y} = \|\boldsymbol{x}\| \|\boldsymbol{y}\| \cos(\theta).$$

This will be important to us only in the case when the angle between $\boldsymbol{x}$ and $\boldsymbol{y}$ is a right angle, and so the cosine is 0.

Definition. Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be in $\mathbb{R}^n$. The $\boldsymbol{x}$ and $\boldsymbol{y}$ are ***orthogonal*** (or perpendicular) if $\boldsymbol{x} \cdot \boldsymbol{y} = 0$. □

Example. The vectors $\boldsymbol{x} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ and $\boldsymbol{y} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$ are orthogonal, since

$$\boldsymbol{x} \cdot \boldsymbol{y} = 2 \cdot 3 + 3 \cdot (-2) = 0.$$

☑