Multivariable Calculus

August 15, 2018

Contents

Vectors

In this section, we review some basics about vectors. This includes the definition of a vector, basic vector operations, standard basis vectors, and notation.

Vectors

In linear algebra, we often worked with vectors. We begin by recalling the (algebraic) definition of a vector in $\mathbb{R}^n$.

Definition 1. *A* vector *in $\mathbb{R}^n$ is an ordered n-tuple of real numbers. That is, a vector $\vec{v}$ may be written as*

$$\vec{v} = (a_1, a_2, ..., a_n)$$

where $a_1, a_2, ..., a_n$ are real numbers.

We call the numbers a_i the components *or* entries *of the vector. We call n the* dimension *of the vector $\vec{v}$, and say that $\vec{v}$ is n-*dimensional.

We write the vector with an arrow above it, as $\vec{v}$, in order to make the distinction between vectors and *scalars*, which are just real numbers. Some other common notations for vectors are $\mathbf{v}$ and $\hat{v}$. It's important to make this distinction between vectors and scalars, so you should make use of one of these notations for vectors.

Example 1. $\vec{v} = (1, 3)$ *is a vector in* $\mathbb{R}^2$.

$\vec{w} = (-1, 5, 0)$ *is a vector in* $\mathbb{R}^3$.

$\vec{x} = (1, -2, 3)$ *is a vector in* $\mathbb{R}^3$.

$\vec{y} = (-6, \pi, 1/24, -0.5, 3)$ *is a vector in* $\mathbb{R}^5$.

It's sometimes convenient to write a vector as a column vector instead (particularly when working with linear transformations, which we'll review in a later section). We could write

$$\vec{v} = (a_1, a_2, ..., a_n) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

Learning outcomes:
Author(s):

or

$$\vec{v} = (a_1, a_2, ..., a_n) = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

The choice between square brackets and parentheses is just a difference in notation, they mean the same thing, and you should feel free to use either.

Example 2. *We write the following vectors as column vectors.*

$\vec{v} = (1, 3) = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$

$\vec{w} = (-1, 5, 0) = \begin{bmatrix} -1 \\ 5 \\ 0 \end{bmatrix}.$

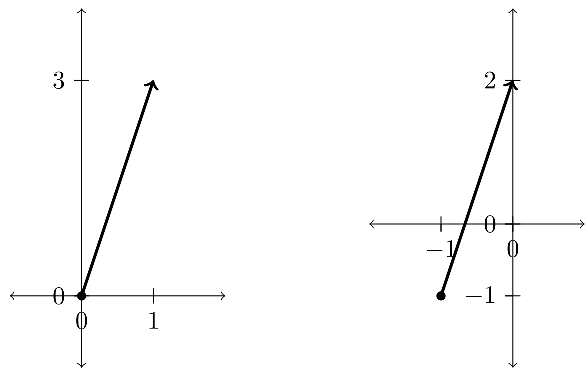
$\vec{x} = (1, -2, 3) = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}.$

$\vec{y} = (-6, \pi, 1/24, -0.5, 3) = \begin{bmatrix} -6 \\ \pi \\ 1/24 \\ -0.5 \\ 3 \end{bmatrix}.$

The Geometric Perspective

We also can think of a vector geometrically, as giving a direction and magnitude, but without a fixed position.

In two or three dimensions, it is useful to visualize a vector as an arrow in $\mathbb{R}^n$. We might visualize a vector $\vec{v} = (1, 3)$ in $\mathbb{R}^2$ as the arrow starting at the origin and ending at the point $(1, 3)$, thus giving a direction and a magnitude. However, we typically don't think of a vector as having a set location. We could also visualize the vector $\vec{v}$ as starting at the point $(-1, -1)$ and ending at the point $(0, 2)$. Note that this arrow would have the same direction and magnitude as the one starting at the origin, thus they represent the same vector.



In four or higher dimensions, visualizing anything becomes very difficult. It can still be useful to think of a vector $(1, 2, 3, 4, 5)$ in $\mathbb{R}^5$ as starting at the origin and ending at the point $(1, 2, 3, 4, 5)$, but you probably won't be able to have a very clear picture of this in your head.

This concept will probably seem more useful once you think about a displacement vector.

Definition 2. *Given points $P_1 = (x_1, ..., x_n)$ and $P_2 = (y_1, ..., y_n)$ in $\mathbb{R}^n$, the* displacement vector *from P_1 to P_2 is*

$$\vec{P_1P_2} = (y_1 - x_1, ..., y_n - x_n).$$

This is the vector that starts at P_1 and ends at P_2.

Notice that the notation $(a_1, ..., a_n)$ that we use for a vector in $\mathbb{R}^n$ is identical to the notation that we'd use for a point in $\mathbb{R}^n$. Since both vectors and points in $\mathbb{R}^n$ are defined as n-tuples of points, they are, in some sense, the same thing. The difference between the two comes when we consider the context and geometric significance of the vector or point that we're working with. As we move into multivariable calculus, we'll often blur the distinction between a vector and a point, and sometimes think of a vector as a point and vice versa. This will be greatly simplify notation, and we promise that it won't be as confusing as it sounds!

Vector Operations

Before defining some basic vector operations, we define what it means for two vectors to be equal. This is done by comparing the components of the vectors.

Definition 3. *Two vectors $(a_1, a_2, ..., a_n)$ and $(b_1, b_2, ..., b_n)$ in $\mathbb{R}^n$ are* equal *if their corresponding components are equal, so $a_1 = b_1$, $a_2 = b_2$, ... , $a_n = b_n$.*

Notice that, in order to be equal, two vectors must have the same dimension and the same entries in the same order. Thus, the vectors $(1,3)$ and $(1,3,0)$ are not equal.

We now define addition of two vectors of the same dimension, which is done componentwise.

Definition 4. *Let $\vec{a} = (a_1, a_2, ..., a_n)$ and $\vec{b} = (b_1, b_2, ..., b_n)$ be vectors in $\mathbb{R}^n$. We define $\vec{a} + \vec{b}$ to be the vector in $\mathbb{R}^n$ given by*

$$\vec{a} + \vec{b} = (a_1 + b_1, a_2 + b_2, ..., a_n + b_n).$$

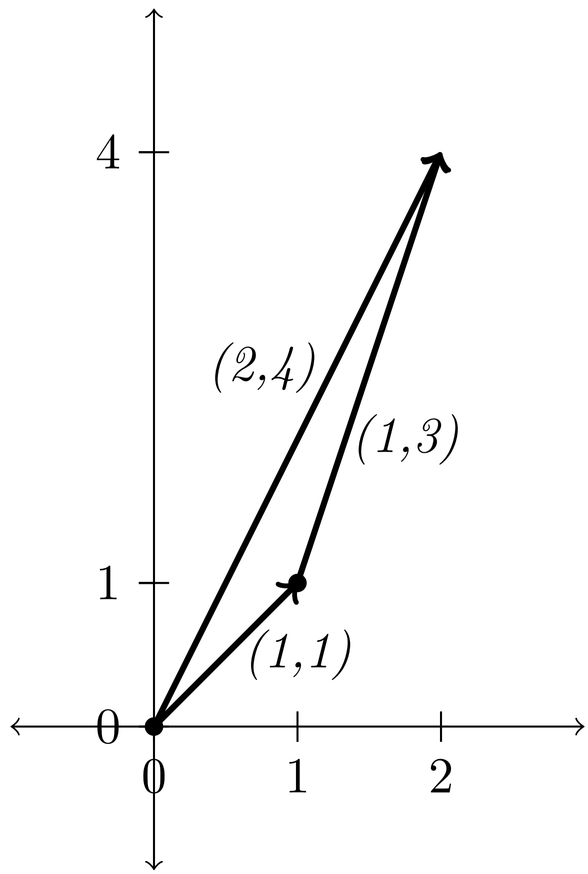
Note that we can only add two vectors if they have the same dimension.

Example 3. *Adding the vectors $\vec{a} = (1, -8, 2)$ and $\vec{b} = (3, -1, -2)$, we obtain*

$$\begin{aligned}\vec{a} + \vec{b} &= (1, -8, 2) + (3, -1, -2) \\ &= (1 + 3, -8 - 1, 2 - 2) \\ &= (4, -9, 0).\end{aligned}$$

Geometrically, we can add vectors by placing the start point of the second vector at the end point of the first vector, and drawing an arrow from the start point of the first vector to the end point of the second vector.

Example 4. *In this example, we add the vectors $(1,1)$ and $(1,3)$. Adding these vectors algebraically, we obtain $(2,4)$. We can also see this geometrically by placing the start point of the vector $(1,3)$ at the end of the vector $(1,1)$ (so at the point $(1,1)$), and drawing the vector from the origin to the end point of the vector $(1,3)$, which is now at $(2,4)$.*



Another vector operation is scalar multiplication. Here, we multiply a vector by a real number, possibly changing the length of the vector.

Definition 5. *Let $\vec{a} = (a_1, ..., a_n)$ be a vector in $\mathbb{R}^n$, and let r be a real number (also called a scalar). We define the* scalar product *$r\vec{a}$ to be*

$$r\vec{a} = (ra_1, ..., ra_n).$$

Thus, we see that scalar multiplication is defined by multiplying each component of the vector by the scalar r.

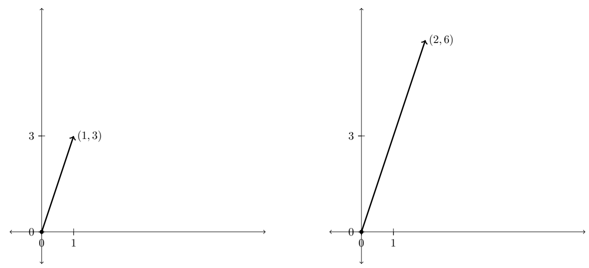
Example 5. $3(1, 5, -2) = (3, 15, -6)$

$-1(1, 1, 1) = (-1, -1, -1)$

$0(6, 2, 4) = (0, 0, 0)$

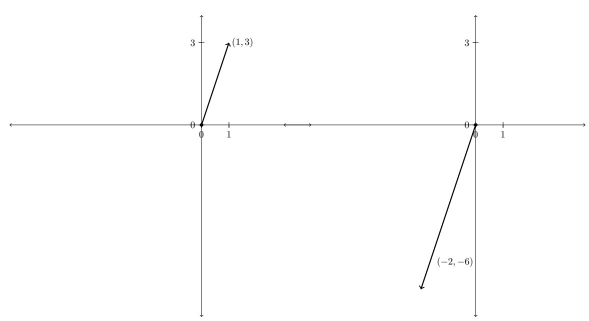
Now, let's look at what scalar multiplication does geometrically. Consider the vector $(1, 3)$.

When we multiply $(1,3)$ by 2, we obtain $(2,6)$, which is twice as long as $(1,3)$ and goes in the same direction.



When we multiply $(1,3)$ by $\dfrac{1}{2}$, we obtain $\left(\dfrac{1}{2}, \dfrac{3}{2}\right)$, which is half as long as $(1,3)$ and goes in the same direction.

If we multiply $(1,3)$ by -2, we obtain $(-2,-6)$, which is twice as long as $(1,3)$ and goes in the exact opposite direction.



Thus, we have seen that multiplying a vector by a scalar changes the length of a vector, but not the direction (except for reversing it, if the scalar is negative).

Properties

Now, let's recall some useful properties of vector addition and scalar multiplication.

Proposition 1. *Suppose $\vec{a}, \vec{b}, \vec{c}$ are vectors in $\mathbb{R}^n$ and k, l are real numbers. Then*

(a) $\vec{a} + \vec{b} = \vec{b} + \vec{a}$ *(vector addition is* commutative*);*

(b) $\vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c}$ *(vector addition is* associative*);*

(c) $\vec{a} + \vec{0} = \vec{0} + \vec{a} = \vec{a}$*, where* $\vec{0} = (0, ..., 0)$ *is the* zero vector *in* $\mathbb{R}^n$*;*

(d) $(k + l)\vec{a} = k\vec{a} + l\vec{a}$*;*

(e) $k(\vec{a} + \vec{b}) = k\vec{a} + k\vec{b}$ *(with the previous property, scalar multiplication is* distributive*);*

(f) $k(l\vec{a}) = (kl)\vec{a} = l(k\vec{a})$*;*

(g) $1\vec{a} = \vec{a}$*.*

These properties tell us different kinds of algebraic manipulations that we can do with vectors.

Standard Basis Vectors

It's often useful to write things in terms of the standard basis vectors for $\mathbb{R}^n$.

Definition 6. *The vectors* $\vec{e}_1 = (1, 0, ...0)$*,* $\vec{e}_2 = (0, 1, 0, ..., 0)$*, ...,* $\vec{e}_n = (0, ..., 0, 1)$ *in* $\mathbb{R}^n$ *are called the* standard basis vectors *for* $\mathbb{R}^n$*.*

Note that any vector in $\mathbb{R}^n$ can be written uniquely as a linear combination of the standard unit vectors. For example, in $\mathbb{R}^4$,

$$\begin{aligned}(1, 5, -3, 6) &= 1(1, 0, 0, 0) + 5(0, 1, 0, 0) - 3(0, 0, 1, 0) + 6(0, 0, 0, 1)\\ &= 1\vec{e}_1 + 5\vec{e}_2 - 3\vec{e}_3 + 6\vec{e}_4.\end{aligned}$$

In $\mathbb{R}^2$, we sometimes write the standard basis vectors as $\mathbf{i} = (1, 0)$ and $\mathbf{j} = (0, 1)$. This gives us a new notation for vectors, for example we could write

$$(3, 4) = 3\mathbf{i} + 4\mathbf{j}.$$

Similarly, in $\mathbb{R}^3$, we sometimes write the standard basis vectors as $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, and $\mathbf{k} = (0, 0, 1)$. We can then write

$$(2, 3, 4) = 2\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}.$$

Summary

In this section, we reviewed some basics about vectors, including the definition of a vector, basic vector operations, standard basis vectors, notation, and the geometric perspective.

The Dot Product

In this section we review the dot product on vectors. This also includes the angle between vectors and the projection of one vector onto another.

The Dot Product

We begin with the definition of the dot product.

Definition 7. *The* dot product *of two vectors* $\vec{v} = (v_1, v_2, ..., v_n)$ *and* $\vec{w} = (w_1, w_2, ..., w_n)$ *in* $\mathbb{R}^n$ *is*

$$\vec{v} \cdot \vec{w} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n.$$

Notice that the dot product takes two vectors and outputs a scalar.

Example 6. $(1, 6) \cdot (-3, -6) = -3 - 36 = -39$

$(1, 2, 3) \cdot (7, -2, 4) = 7 - 4 + 12 = 15$

$(1, 7, -3) \cdot (3, 0, 1) = 3 + 0 - 3 = 0$

We can also compute the dot product using the magnitude (or length) of the vectors and the angle in between them.

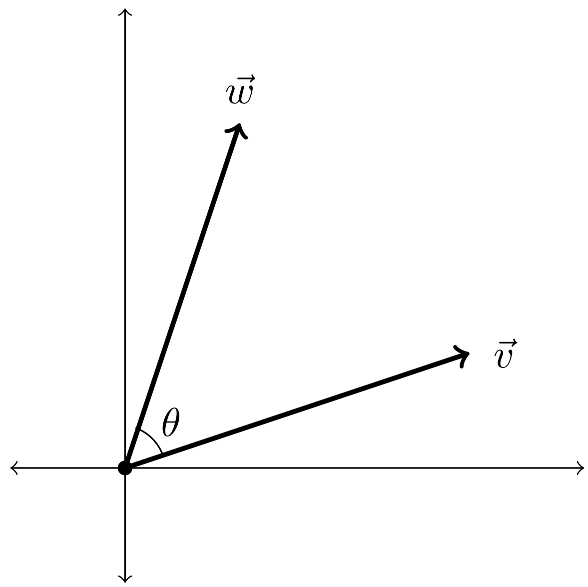
Proposition 2. *If* $\vec{v}$ *and* $\vec{w}$ *are vectors in* $\mathbb{R}^n$*, then*

$$\vec{v} \cdot \vec{w} = \|\vec{v}\| \, \|\vec{w}\| \cos\theta,$$

where $\|\vec{v}\|$ *and* $\|\vec{w}\|$ *are the lengths of the vectors* $\vec{v}$ *and* $\vec{w}$*, respectively, and* θ *is the angle between* $\vec{v}$ *and* $\vec{w}$*.*

This is illustrated in the picture below.

Learning outcomes:
Author(s):



This provides us with a geometric interpretation of the dot product: it gives us a measure of "how much" in the same direction two vectors are (taking their lengths into account). This also gives us a useful way to compute the angle between two vectors.

Example 7. *Consider the vectors* $(1,4)$ *and* $(-2,2)$*. We have*

$$\begin{aligned}(1,4)\cdot(-2,2) &= -2+8=6,\\ \|(1,4)\| &= \sqrt{1^2+4^2}=\sqrt{17},\\ \|(-2,2)\| &= \sqrt{(-2)^2+2^2}=\sqrt{8}.\end{aligned}$$

From $\vec{v}\cdot\vec{w}=\|\vec{v}\|\,\|\vec{w}\|\cos\theta$*, we then have*

$$6=\sqrt{17}\sqrt{8}\cos\theta.$$

Solving for θ*, we obtain the angle between the vectors as*

$$\theta=\arccos\left(\frac{6}{\sqrt{17}\sqrt{8}}\right)\approx 59.04^\circ$$

Furthermore, note that for nonzero vectors $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^n$, their dot product is 0 if and only if $\cos(\theta)=0$. This means that θ would have to be 90° or 270°, meaning that $\vec{v}$ and $\vec{w}$ are perpendicular.

Proposition 3. *Two nonzero vectors* $\vec{v}$ *in* $\vec{w}$ *in* $\mathbb{R}^n$ *are perpendicular if and only if* $\vec{v}\cdot\vec{w}=0$.

This provides us with a very useful algebraic method for determining if two vectors are perpendicular.

Example 8. *The vectors* $(1, 7, -3)$ *and* $(3, 0, 1)$ *in* $\mathbb{R}^3$ *are perpendicular, since*

$$(1, 7, -3) \cdot (3, 0, 1) = 3 + 0 - 3 = 0.$$

By taking the dot product of a vector with itself, we get an important relationship between the dot product and the length of a vector.

Proposition 4. *Let* $\vec{v}$ *be a vector in* $\mathbb{R}^n$*. Then*

$$\vec{v} \cdot \vec{v} = \|\vec{v}\|^2.$$

This can be shown directly, or using the fact that the angle between $\vec{v}$ and itself is 0.

Projection of one vector onto another

We can also use the dot product to define the projection of one vector onto another.

Definition 8. *For vectors* $\vec{a}$ *and* $\vec{b}$ *in* $\mathbb{R}^n$*, we define the* vector projection *of* $\vec{a}$ *onto* $\vec{b}$ *as*

$$proj_{\vec{b}}(\vec{a}) = \frac{\vec{a} \cdot \vec{b}}{\|\vec{b}\|^2}\vec{b} = \frac{\vec{a} \cdot \vec{b}}{\vec{b} \cdot \vec{b}}\vec{b}.$$

Example 9. *We can use this to find the projection of* $(2, 4, 3)$ *onto* $(1, -1, 1)$*.*

$$\begin{aligned} proj_{(1,-1,1)}(2, 4, 3) &= \frac{(2, 4, 3) \cdot (1, -1, 1)}{(1, -1, 1) \cdot (1, -1, 1)}(1, -1, 1) \\ &= \frac{2 - 4 + 3}{1 + 1 + 1}(1, -1, 1) \\ &= \frac{1}{3}(1, -1, 1) \\ &= \left(\frac{1}{3}, -\frac{1}{3}, \frac{1}{3}\right) \end{aligned}$$

Summary

In this section we reviewed the dot product on vectors, the angle between vectors, and the projection of one vector onto another.

The Cross Product

In this section, we review the vector cross product, including the geometric perspective of the cross product, the area of a parallelogram, and the volume of parallelepiped.

The Cross Product

The cross product is fundamentally different from the dot product in a couple of ways. The cross product is defined only on vectors in $\mathbb{R}^3$, while the dot product is defined in $\mathbb{R}^n$ for any positive integer n. Furthermore, the cross product takes two vectors and produces another vector, while the dot product takes two vectors and produces a scalar.

We now give the algebraic definition of the cross product.

Definition 9. *Let $\vec{a} = (a_1, a_2, a_3)$ and $\vec{b} = (b_1, b_2, b_3)$ be vectors in $\mathbb{R}^3$. The* cross product *of $\vec{a}$ and $\vec{b}$, denoted $\vec{a} \times \vec{b}$, is defined to be*

$$\vec{a} \times \vec{b} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}.$$

Equivalently, we can compute the cross product as

$$\vec{a} \times \vec{b} = det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix},$$

where

$$\begin{aligned} \mathbf{i} &= (1, 0, 0), \\ \mathbf{j} &= (0, 1, 0), \\ \mathbf{k} &= (0, 0, 1). \end{aligned}$$

Learning outcomes:
Author(s):

Example 10.

$$
\begin{aligned}
(3,2,-1)\times(9,0,2) &= det\begin{pmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 3 & 2 & -1 \\ 9 & 0 & 2 \end{pmatrix} \\
&= (2\cdot 2)\mathbf{i} - (0\cdot -1)\mathbf{i} + (-1\cdot 9)\mathbf{j} - (2\cdot 3)\mathbf{j} + (3\cdot 0)\mathbf{k} - (9\cdot -1)\mathbf{k} \\
&= 4\mathbf{i} - 15\mathbf{j} + 9\mathbf{k} \\
&= (4,-15,9)
\end{aligned}
$$

The cross product has some nice algebraic properties, which can be very useful.

Proposition 5. *Let $\vec{a}$, $\vec{b}$, and $\vec{c}$ be vectors in $\mathbb{R}^3$, and let $k \in \mathbb{R}$ be a scalar. The cross product has the following properties:*

(a) $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$ *(the cross product is* anticommutative*);*

(b) $\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}$*;*

(c) $(\vec{a} + \vec{b}) \times \vec{c} = \vec{a} \times \vec{c} + \vec{b} \times \vec{c}$ *(with the previous property, the cross product is* distributive *over vector addition);*

(d) $k(\vec{a} \times \vec{b}) = (k\vec{a}) \times \vec{b} = \vec{a} \times (k\vec{b})$.

In particular, it's important to remember that the cross product is *not* commutative, so the order of the vectors matters!

Geometry of the Cross Product

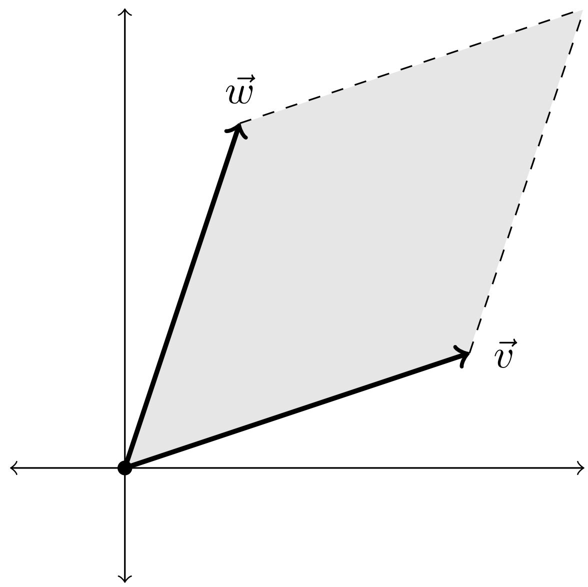
It's often easiest to compute cross products algebraically, but it's easier to understand their significance from a geometric perspective. We now discuss some of the geometric properties of the cross product.

Proposition 6. *Let $\vec{a}$ and $\vec{b}$ be vectors in $\mathbb{R}^3$, and consider their cross product $\vec{a} \times \vec{b}$.*

- *The* magnitude *of the vector $\vec{a} \times \vec{b}$ can be computed as*

$$\|\vec{a} \times \vec{b}\| = \|\vec{a}\|\,\|\vec{b}\|\sin(\theta),$$

where θ is the angle between the vectors $\vec{a}$ and $\vec{b}$. Furthermore, this magnitude is equal to the area of the parallelogram determined by $\vec{a}$ and $\vec{b}$.



- *The vector $\vec{a} \times \vec{b}$ is always perpendicular to both $\vec{a}$ and $\vec{b}$, and follows the* right-hand rule*. That is, if you take your right hand and orient it so you can curl your fingers from the vector $\vec{a}$ to the $\vec{b}$, your thumb will be pointing in the same direction as the cross product $\vec{a} \times \vec{b}$.*

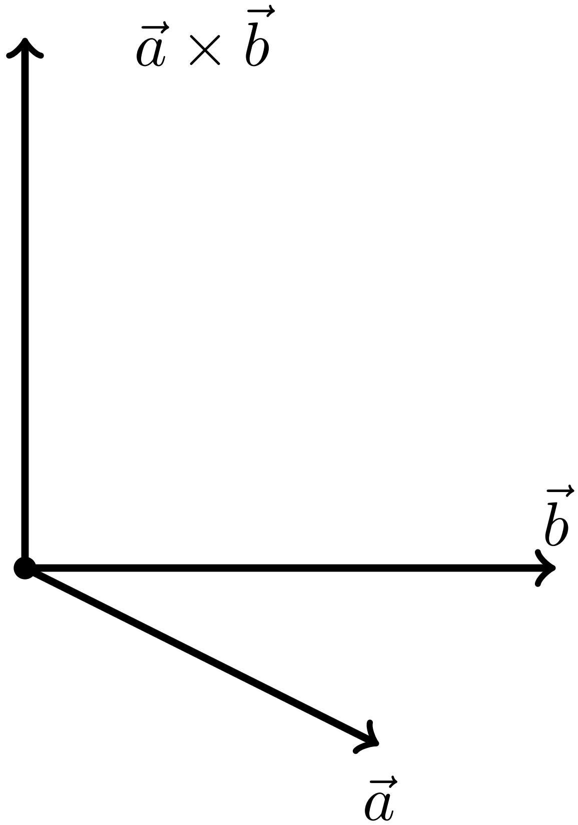
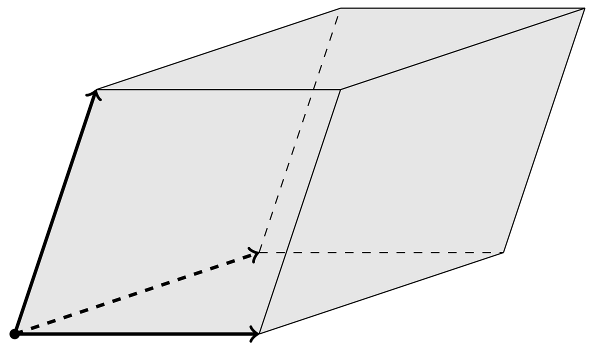


Image this image in $\mathbb{R}^3$, so that $\vec{a} \times \vec{b}$ is perpendicular to both $\vec{a}$ and $\vec{b}$.

Volume of a Parallelepiped

We can use the cross product and dot product together to compute the volume of a parallelepiped.



The volume of the parellelepiped can be computed as the area of the base times the height. We've seen that the area of the base can be computed as the magnitude of a cross product, $\|\vec{a} \times \vec{b}\|$. The height of the parallelepiped can be computed as $\|\vec{c}\| \, |\cos(\theta)|$, where θ is the angle between the vector $\vec{c}$ and a line perpendicular to the base. We then have that the volume is $\|\vec{a} \times \vec{b}\| \, \|\vec{c}\| \, |\cos(\theta)|$, which we can recognize as the absolute value of the dot product of the vectors $\vec{a} \times \vec{b}$ and $\vec{c}$. Thus we have the following proposition.

Proposition 7. *The volume of the parallelepiped determined by the vectors $\vec{a}$, $\vec{b}$, and $\vec{c}$ can be computed as $|(\vec{a} \times \vec{b}) \cdot \vec{c}|$.*

Summary

We've reviewed the cross product, including its properties and geometric perspective, including its use in finding the area of parallelograms and volume of parallelepipeds.

Matrices

In this section, we review matrices, including the determinant and the linear transformation represented by a matrix.

Matrices

We begin with the definition of a matrix.

Definition 10. *An $m \times n$* matrix *A is a rectangular array of numbers a_{ij}, with m rows and n columns:*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

where the a_{ij} are real numbers for i and j integers with $1 \leq i \leq m$ and $1 \leq j \leq n$. The numbers a_{ij} are called the entries *of the matrix A.*

Note that for an entry a_{ij}, the subscript ij describes the location of a_{ij} in the matrix A: i gives the row, and j gives the column.

We can also think of a matrix as a "vector of vectors" in two different ways. If we imagine that the columns of A are vectors in $\mathbb{R}^n$, then the matrix of A can be viewed as a vector of column vectors. If we imagine that the rows of A are vectors in $\mathbb{R}^n$, then the matrix A can be viewed as a vector of row vectors.

Matrix Operations

Here, we'll define matrix addition and matrix multiplication.

In order to be able to add two matrices, they need to have the exact same dimensions. That is, they both need to be $m \times n$ matrices for some fixed values of m and n. When we have two matrices with the same dimensions, we define their sum component-wise or entry-wise.

Learning outcomes:
Author(s):

Definition 11. *Let A and B be two $m \times n$ matrices, with*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

$$B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}.$$

Then we define the matrix sum *$A + B$ to be*

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn}. \end{pmatrix},$$

That is, $A+B$ is the $m \times n$ matrix obtained by adding the corresponding entries of A and B.

Example 11. *We can add the matrices $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ and $B = \begin{pmatrix} 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix}$ as follows:*

$$\begin{aligned} A + B &= \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix} \\ &= \begin{pmatrix} 1+7 & 2+8 & 3+9 \\ 4+10 & 5+11 & 6+12 \end{pmatrix} \\ &= \begin{pmatrix} 8 & 10 & 12 \\ 14 & 16 & 18 \end{pmatrix} \end{aligned}$$

Example 12. *We cannot add the matrices $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ and $B = \begin{pmatrix} 7 & 8 \\ 10 & 11 \end{pmatrix}$, because their dimensions don't match.*

As you might expect, matrix addition has some nice properties which are inherited from addition of real numbers. We list some of them here.

Proposition 8. *Let A, B, and C be $m \times n$ matrices. Then we have*

(a) $A + B = B + A$ *(matrix addition is* commutative*);*

(b) $A + (B + C) = (A + B) + C$ *(matrix addition is* associative*).*

Furthermore, there is an $m \times n$ matrix O, called the zero matrix, *such that $A + O = A$ for any $m \times n$ matrix A. All of the entries of the zero matrix are the real number 0.*

We've seen that matrix addition works in a very natural way, and multiplying a matrix by a scalar (or real number) is similarly nice. We now define scalar multiplication for matrices.

Definition 12. *Let*

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

be an $m \times n$ matrix, and let $k \in \mathbb{R}$ be a scalar. Then the scalar product *of k and A, denoted kA, is*

$$kA = \begin{pmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & & \vdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{pmatrix}.$$

That is, we obtain the scalar product by multiplying each entry in A by the scalar k.

Example 13. *We can compute the scalar product of* 2 *and the matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ as follows:*

$$\begin{aligned} 2A &= 2\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \\ &= \begin{pmatrix} 2\cdot 1 & 2\cdot 2 & 2\cdot 3 \\ 2\cdot 4 & 2\cdot 5 & 2\cdot 6 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{pmatrix}. \end{aligned}$$

We now list some nice properties of scalar multiplication.

Proposition 9. *Let A and B be $m \times n$ matrices, and let k and l be scalars in $\mathbb{R}$. Then*

(a) $(k+l)A = kA + lA$ *(scalar multiplication is* distributive *over scalar addition);*

(b) $k(A+B) = kA + kB$ *(scalar multiplication is* distributive *over matrix addition);*

(c) $k(lA) = (kl)A = l(kA)$.

We'll now define matrix multiplication, which can be a bit trickier to work with than matrix addition or scalar multiplication. Here are some important things to remember about matrix multiplication:

- Not all matrices can be multiplied. In order to compute the product AB of two matrices A and B, the number of columns in A needs to be the same as the number of rows in B.
- Matrix multiplication is *not* commutative. In fact, its possible that the matrix product AB exists but the product BA does not.

Definition 13. *Let A be an $m \times n$ matrix, and let B be an $n \times p$ matrix, with*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

$$B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix}.$$

Note that we are assuming the number of columns in A is the same as the number of rows in B.

We define the matrix product *of A and B, denoted AB, to be the $m \times p$ matrix*

$$AB = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} & a_{11}b_{12} + a_{12}b_{22} + \cdots + a_{1n}b_{n2} & \cdots & a_{11}b_{1p} + a_{12}b_{2p} + \cdots + a_{1n}b_{np} \\ a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1} & a_{21}b_{12} + a_{22}b_{22} + \cdots + a_{2n}b_{n2} & \cdots & a_{21}b_{1p} + a_{22}b_{2p} + \cdots + a_{2n}b_{np} \\ \vdots & \vdots & & \vdots \\ a_{m1}b_{11} + a_{m2}b_{21} + \cdots + a_{mn}b_{n1} & a_{m1}b_{12} + a_{m2}b_{22} + \cdots + a_{mn}b_{n2} & \cdots & a_{m1}b_{1p} + a_{m2}b_{2p} + \cdots + a_{mn}b_{np} \end{pmatrix}$$

Equivalently, we could define the ijth entry of AB to be the dot product of the ith row of A with the jth column of B. This makes sense, since the number of columns in A is the same as the number of rows in B (both n), which ensures that the ith row of A and the jth column of B are both vectors in $\mathbb{R}^n$.

This definition can seem a bit convoluted, and it's easier to understand how matrix multiplication works by going through an example.

Example 14. *We can compute the product AB of the matrices $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$*

and $B = \begin{pmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{pmatrix}$ *as follows:*

$$\begin{aligned} AB &= \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{pmatrix}, \\ &= \begin{pmatrix} 1\cdot 7 + 2\cdot 9 + 3\cdot 11 & 1\cdot 8 + 2\cdot 10 + 3\cdot 12 \\ 4\cdot 7 + 5\cdot 9 + 6\cdot 11 & 4\cdot 8 + 5\cdot 10 + 6\cdot 12 \end{pmatrix}, \\ &= \begin{pmatrix} 58 & 64 \\ 139 & 154 \end{pmatrix}. \end{aligned}$$

Note that we multiplied a 2×3 *matrix by a* 3×2 *matrix, and we obtained a* 2×2 *matrix.*

We can also compute the product BA *for the same matrices as above.*

$$\begin{aligned} BA &= \begin{pmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \\ &= \begin{pmatrix} 7\cdot 1 + 8\cdot 4 & 7\cdot 2 + 8\cdot 5 & 7\cdot 3 + 8\cdot 6 \\ 9\cdot 1 + 10\cdot 4 & 9\cdot 2 + 10\cdot 5 & 9\cdot 3 + 10\cdot 6 \\ 11\cdot 1 + 12\cdot 4 & 11\cdot 2 + 12\cdot 5 & 11\cdot 3 + 12\cdot 6 \end{pmatrix}, \\ &= \begin{pmatrix} 39 & 54 & 69 \\ 49 & 68 & 81 \\ 59 & 82 & 105 \end{pmatrix}. \end{aligned}$$

Note that we multiplied a 3×2 *matrix by a* 2×3 *matrix, and we obtained a* 3×3 *matrix.*

Note that in this case $AB \neq BA$*; matrix multiplication is not commutative, so the order of the matrices matters!*

Although matrix multiplication is not commutative, it still has some nice algebraic properties. We list some of them here.

Proposition 10. *Let* A*,* B*, and* C *be matrices of dimensions such that the following operations are defined, and let* k *be a scalar. Then*

(a) $A(BC) = (AB)C$ *(matrix multiplication is* associative*);*

(b) $k(AB) = (kA)B = A(kB)$*;*

(c) $A(B + C) = AB + AC$*;*

(d) $(A + B)C = AC + BC$ *(with the previous property, matrix multiplication is* distributive *over matrix addition).*

Determinants

When we have a square matrix (meaning an $n \times n$ matrix, where the number of rows and number of columns are the same), we can compute an important number, called the determinant of the matrix. It turns out that this single number can tell us some important things about the matrix!

We begin by defining the determinant of a 2×2 matrix.

Definition 14. *Consider the 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. We define the* determinant *of the matrix A to be*

$$det(A) = ad - bc.$$

We also sometimes use the notation $|A|$ for the determinant of the matrix A.

Note that the determinate of a 2×2 matrix is just a number, not a matrix. We compute the determinant in a couple of examples.

Example 15. *We'll compute the determinant of the matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.*

$$\begin{aligned} det(A) &= 1 \cdot 4 - 2 \cdot 3 \\ &= -2. \end{aligned}$$

We've defined the determinant of 2×2 matrices, but we haven't defined the determinant of a larger square matrix yet. It turns out that the determinant is defined *inductively*. This means that the determinant of a 3×3 matrix is defined using determinants of 2×2 matrices, the determinant of a 4×4 matrix is defined using determinants of 3×3 matrices, the determinant of a 5×5 matrix is defined using determinants of 4×4 matrices, and so on. This means in order to compute the determinant of a large square matrix, we often need to compute the determinants of many smaller matrices.

We now give the definition of the determinant of an $n \times n$ matrix.

Definition 15. *Let A be an $n \times n$ matrix, with entries a_{ij}. We defined the* determinant *of A to be the number computed by*

$$det(A) = (-1)^{1+1} a_{11}\, det(A_{11}) + (-1)^{1+2} a_{12}\, det(A_{12}) + \cdots + (-1)^{1+n} a_{1n}\, det(A_{1n}),$$

where A_{ij} is the $(n-1) \times (n-1)$ submatrix of A which we obtain by deleting the ith row and $j$the column from A.

This definition is pretty confusing if you read through it without seeing an example, but this actually follows a nice pattern. This pattern is easier to see with an example.

Example 16. *We compute the determinant of the 4×4 matrix,*

$$A = \begin{pmatrix} 1 & 4 & 2 & -1 \\ 0 & 0 & -2 & 1 \\ -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Note that we begin by writing this in terms of determinants of 3×3 matrices. But in order to compute the determinant of each 3×3 matrix, we write it in terms of 2×2 matrices! This winds up being a lot of determinants to compute.

$$\begin{aligned} det(A) = (-1)^{1+1} 1 det \begin{pmatrix} 0 & -2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + (-1)^{1+2} 4 det \begin{pmatrix} 0 & -2 & 1 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ + (-1)^{1+3} 2 det \begin{pmatrix} 0 & 0 & 1 \\ -3 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + (-1)^{1+4} (-1) det \begin{pmatrix} 0 & 0 & -2 \\ -3 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

We now compute the determinant of each of the 3×3 submatrices, which we *compute using determinants of* 2×2 *matrices.*

$$
\begin{aligned}
det\begin{pmatrix} 0 & -2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} &= (-1)^{1+1} 0 det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (-1)^{1+2}(-2) det\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\
&\quad + (-1)^{1+3} 1 det\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\
&= 1 \cdot 0 \cdot (1 \cdot 1 - 0 \cdot 0) + -1 \cdot (-2) \cdot (0 \cdot 1 - 0 \cdot 0) + 1 \cdot 1 \cdot (0 \cdot 0 - 1 \cdot 0) \\
&= 0 \\
det\begin{pmatrix} 0 & -2 & 1 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} &= (-1)^{1+1} 0 det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (-1)^{1+2}(-2) det\begin{pmatrix} -3 & 0 \\ 0 & 1 \end{pmatrix} \\
&\quad + (-1)^{1+3} 1 det\begin{pmatrix} -3 & 1 \\ 0 & 0 \end{pmatrix} \\
&= 1 \cdot 0 \cdot (1 \cdot 1 - 0 \cdot 0) + -1 \cdot (-2) \cdot (-3 \cdot 1 - 0 \cdot 0) + 1 \cdot 1 \cdot (-3 \cdot 0 - 1 \cdot 0) \\
&= 6 \\
det\begin{pmatrix} 0 & 0 & 1 \\ -3 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} &= (-1)^{1+1} 0 det\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + (-1)^{1+2} 0 det\begin{pmatrix} -3 & 0 \\ 0 & 1 \end{pmatrix} \\
&\quad + (-1)^{1+3} 1 det\begin{pmatrix} -3 & 0 \\ 0 & 0 \end{pmatrix} \\
&= 1 \cdot 0 \cdot (0 \cdot 1 - 0 \cdot 0) + -1 \cdot 0 \cdot (-3 \cdot 1 - 0 \cdot 0) + 1 \cdot 1 \cdot (-3 \cdot 0 - 0 \cdot 0) \\
&= 0 \\
det\begin{pmatrix} 0 & 0 & -2 \\ -3 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} &= (-1)^{1+1} 0 det\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + (-1)^{1+2} 0 det\begin{pmatrix} -3 & 1 \\ 0 & 0 \end{pmatrix} \\
&\quad + (-1)^{1+3}(-2) det\begin{pmatrix} -3 & 0 \\ 0 & 0 \end{pmatrix} \\
&= 1 \cdot 0 \cdot (0 \cdot 0 - 1 \cdot 0) + -1 \cdot 0 \cdot (-3 \cdot 0 - 1 \cdot 0) + 1 \cdot (-2) \cdot (-3 \cdot 0 - 0 \cdot 0) \\
&= 0
\end{aligned}
$$

Substituting these in to our computation of the determinant of A*, we then obtain*

$$
\begin{aligned}
det(A) &= 1 \cdot 1 \cdot 0 + (-1) \cdot 4 \cdot (6) + 1 \cdot 2 \cdot 0 + (-1) \cdot (-1) \cdot 0 \\
&= -24.
\end{aligned}
$$

We sometimes call this method of computing a determinant as "expanding along the first row." This is because we can also compute the determinant of a matrix by similarly expanding along a different row, or even a column.

Proposition 11. *We can similarly compute the determinant of an* $n \times n$ *matrix*

A by expanding along any row or column. Expanding along the ith row, we have

$$det(A) = (-1)^{i+1} a_{i1}\, det(A_{i1}) + (-1)^{i+2} a_{i2}\, det(A_{i2}) + \cdots + (-1)^{i+n} a_{in}\, det(A_{in}).$$

Expanding along the jth column, we have

$$det(A) = (-1)^{1+j} a_{1j}\, det(A_{1j}) + (-1)^{2+j} a_{2j}\, det(A_{2j}) + \cdots + (-1)^{n+j} a_{nj}\, det(A_{nj}).$$

Once again, A_{ij} is the $(n-1)\times(n-1)$ submatrix obtained from A by deleting the ith row and jth column.

It can be useful to think about which row or column will be easiest to expand along. In particular, choosing a row or column with a lot of zeros greatly simplifies computation.

Example 17. *We'll once again compute the determinant of the 4×4 matrix*

$$A = \begin{pmatrix} 1 & 4 & 2 & -1 \\ 0 & 0 & -2 & 1 \\ -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

this time by expanding along the second column. Note that this column is a good choice, since there's only one nonzero element. We have

$$det(A) = (-1)^{1+2}(4)\, det \begin{pmatrix} 0 & -2 & 1 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + (-1)^{2+2}(0)\, det \begin{pmatrix} 1 & 2 & -1 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + (-1)^{3+2}(0)\, det \begin{pmatrix} 1 & 2 & -1 \\ 0 & -2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

We'll only compute the determinant of the submatrix $\begin{pmatrix} 0 & -2 & 1 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$; we won't bother computing the others since their determinants will be multiplied by 0.

$$\begin{aligned} det \begin{pmatrix} 0 & -2 & 1 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} &= (-1)^{3+1}(0)\, det \begin{pmatrix} -2 & 1 \\ 1 & 0 \end{pmatrix} + (-1)^{3+2}(0)\, det \begin{pmatrix} 0 & 1 \\ -3 & 0 \end{pmatrix} + (-1)^{3+3}(1)\, det \begin{pmatrix} 0 & -2 \\ -3 & 1 \end{pmatrix} \\ &= 0 + 0 + (1)(1)(0\cdot 1 - (-2)\cdot(-3)), \\ &= -6. \end{aligned}$$

Once again, we don't bother computing the determinants which will be multiplied by zero. Note that we chose to expand across the last row, since it had two zeroes. Expanding along the first column would also have been a reasonable choice.

Returning to our computation of the determinant of A, we have

$$\begin{aligned} det(A) &= (-1)^{1+2}(4)\, det \begin{pmatrix} 0 & -2 & 1 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + (-1)^{2+2}(0)\, det \begin{pmatrix} 1 & 2 & -1 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + (-1)^{3+2}(0)\, det \begin{pmatrix} 1 & 2 & -1 \\ 0 & -2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ &= (-1)(4)(-6) + 0 + 0 + 0, \\ &= 24. \end{aligned}$$

Notice that this matching our previous computation, expanding along the first row.

One of the most powerful uses of the determinant is to tell us whether or not a matrix is invertible. Recall that an $n \times n$ matrix A is *invertible* if there is a matrix B such that $AB = BA = I_n$, where I_n is the $n \times n$ identity matrix.

Proposition 12. *An $n \times n$ matrix A is invertible if and only if its determinant is nonzero.*

This gives us a convenient way to test if a matrix is invertible, without needing to produce an explicit inverse.

Example 18. *We found that the determinant of the matrix*

$$A = \begin{pmatrix} 1 & 4 & 2 & -1 \\ 0 & 0 & -2 & 1 \\ -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

is 24*. Since this is nonzero, the matrix A is invertible.*

On the other hand, you can verify that the determinant of the matrix

$$B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & -1 & -2 \\ -3 & 1 & -1 & 0 \\ -1 & 3 & 1 & 2 \end{pmatrix}$$

is 0*. Thus, the matrix B is not invertible.*

Linear Transformations

One of the most important uses of matrices is to represent linear transformations. Recall the definition of a linear transformation.

Definition 16. *A function T from $\mathbb{R}^n$ to $\mathbb{R}^n$ is a linear transformation if for all vectors $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^n$ and all scalars $k \in \mathbb{R}$, we have*

(a) $T(\vec{v} + \vec{w}) = T(\vec{v}) + T(\vec{w})$*;*

(b) $T(k\vec{v}) = kT(\vec{v})$*.*

We can view an $m \times n$ matrix A as representing a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ as follows. We write vectors as column vectors, or, equivalently, $n \times 1$ or $m \times 1$ matrices. For an input column vector $\vec{v}$ in $\mathbb{R}^n$, we multiply $\vec{v}$ by A

on the left, using matrix multiplication. This produces an $m \times 1$ matrix, or, equivalently, a column vector in $\mathbb{R}^m$. Thus, we can define a function

$$T_A(\vec{v}) = A\vec{v}.$$

Using properties of matrix multiplication, we have that this is a linear transformation. Thus, we have the linear transformation associated to a matrix.

Example 19. *Consider the linear transformation T_A from $\mathbb{R}^3$ to $\mathbb{R}^2$ corresponding to the 2×3 matrix*

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

Let investigate the images of several vectors in $\mathbb{R}^3$ under the linear transformation T_A.

$$\begin{aligned} T_A((1,2,3)) &= \begin{pmatrix} 1 & -5 & 3 \\ 2 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \\ &= \begin{pmatrix} 1 \cdot 1 + -5 \cdot 2 + 3 \cdot 3 \\ 2 \cdot 1 + 0 \cdot 2 + -1 \cdot 3 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ -1 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} T_A((1,-1,2)) &= \begin{pmatrix} 1 & -5 & 3 \\ 2 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 \cdot 1 + -5 \cdot -1 + 3 \cdot 1 \\ 2 \cdot 1 + 0 \cdot -1 + -1 \cdot 1 \end{pmatrix} \\ &= \begin{pmatrix} 9 \\ 1 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} T_A((1,0,0)) &= \begin{pmatrix} 1 & -5 & 3 \\ 2 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 \cdot 1 + -5 \cdot 0 + 3 \cdot 0 \\ 2 \cdot 1 + 0 \cdot 0 + -1 \cdot 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 2 \end{pmatrix} \end{aligned}$$

$$\begin{aligned}
T_A((0,1,0)) &= \begin{pmatrix} 1 & -5 & 3 \\ 2 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 \cdot 0 + -5 \cdot 1 + 3 \cdot 0 \\ 2 \cdot 0 + 0 \cdot 1 + -1 \cdot 0 \end{pmatrix} \\
&= \begin{pmatrix} -5 \\ 0 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
T_A((0,0,1)) &= \begin{pmatrix} 1 & -5 & 3 \\ 2 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 \cdot 0 + -5 \cdot 0 + 3 \cdot 1 \\ 2 \cdot 0 + 0 \cdot 0 + -1 \cdot 1 \end{pmatrix} \\
&= \begin{pmatrix} 3 \\ -1 \end{pmatrix}
\end{aligned}$$

Notice that when we apply the linear transformation to the standard unit vectors $\vec{e}_1$, $\vec{e}_2$, and $\vec{e}_3$, we obtain the columns of A as the output vector. This observation can be used to reconstruct a matrix from a given linear transformation.

Proposition 13. *Given any linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$, there is an $m \times n$ matrix such that $T = T_A$.*

Furthermore, the columns of A can be obtained by applying T to the standard unit vectors. More specifically, the jth column of A is given by $T(\vec{e}_j)$.

We can see how this is useful through an example.

Example 20. *Consider the linear transformation T from $\mathbb{R}^2$ to $\mathbb{R}^2$ that rotates a vector by 30° counterclockwise. We can see geometrically that, for the standard unit vectors $\vec{e}_1$ and $\vec{e}_2$ in $\mathbb{R}^2$, we have*

$$\begin{aligned}
T((1,0)) &= \left(\frac{\sqrt{3}}{2}, \frac{1}{2} \right), \\
T((0,1)) &= \left(-\frac{1}{2}, \frac{\sqrt{3}}{2} \right).
\end{aligned}$$

These tell us the columns of the matrix corresponding to the linear transformation, so we then know that the rotation can be represented by the matrix

$$A = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}.$$

Summary

In this section, we reviewed matrix operations and properties, determinants, and linear transformations.

Although we've reviewed some of the most important concepts from linear algebra, there is still a lot of material that we weren't able to include here. Make sure you refer back to your linear algebra textbook if there's anything else you need to review!

Representations of Lines and Planes

In this section, we review the different ways we can represent lines and planes, including parametric representations.

Representations of Lines

When you think of describing a line algebraically, you might think of the standard form

$$y = mx + b,$$

where m is the slope and b is the y-intercept. This is often called *slope-intercept* form.

In addition to slope-intercept form, there are several other ways to represent lines. For example, you may remember using *point-slope* form in single variable calculus. We can describe a line of slope m going through a point (x_0, y_0) with the equation

$$y - y_0 = m(x - x_0).$$

It's important to note that there are many different possible choices for the point (x_0, y_0). Because of this, unlike slope-intercept form, point-slope form does not give a unique representation of a line.

In linear algebra, we saw that we could parametrize a line using a vector $\vec{v} = (v_1, v_2)$ giving the direction of the line, and a point (x_0, y_0) that the line passes through. We parametrize the line as

$$\begin{aligned} \vec{x}(t) &= (v_1, v_2)t + (x_0, y_0), \\ &= (v_1 t + x_0, v_2 t + y_0). \end{aligned}$$

Note that this representation works a bit differently from the previous two representations. In slope-intercept form and point-slope form, the line was the set of points (x, y) satisfying the given equation. However, in the parametrization, we plug in values for the parameter t in order to get points on the line.

Unlike slope-intercept form and point-slope form, the parametrization of a line can easily be generalized to three or more dimensions. That is, a line in $\mathbb{R}^n$ through the point $\vec{a}$ and in the direction of the vector $\vec{v}$ can be parametrized as

$$\vec{x}(t) = \vec{v}t + \vec{a},$$

Learning outcomes:
Author(s):

for $t \in \mathbb{R}$.

If we would like to describe a line in higher dimensions using equations (rather than a parametrization), we would need more than one equation. For example, in $\mathbb{R}^3$, we would require two equations to determine a line.

Representations of Planes

We also have multiple ways to represent planes. Here, we'll focus on planes in $\mathbb{R}^3$.

Recall that a plane can be determined by two vectors (giving the "direction" of the plane) and a point that the plane passes through. We can use this to give a parametrization for the plane through the point $\vec{a}$ and parallel to the vectors $\vec{v}$ and $\vec{w}$:

$$\vec{x}(s,t) = \vec{v}s + \vec{w}t + \vec{a},$$

for s and t in $\mathbb{R}$. Note that we require two parameters for the parametrization of the plane.

We can also describe a plane using a single linear equation in x, y, and z. For example,

$$2x + 4y - z = 9$$

defines a plane. A standard way to do this is using a point on the plane and a normal vector to the plane. Recall that a normal vector is perpendicular to every vector in the plane. If $\vec{n} = (n_1, n_2, n_3)$ is a normal vector to a plane passing through the point $\vec{a} = (a_1, a_2, a_3)$, the plane is defined by the equation

$$\vec{n} \cdot (\vec{x} - \vec{a}) = 0.$$

This can be rewritten as

$$n_1(x - a_1) + n_2(y - a_2) + n_3(z - a_3) = 0.$$

Summary

We reviewed various representations of lines and planes, including parametrizations.