Contents

Chapter 1

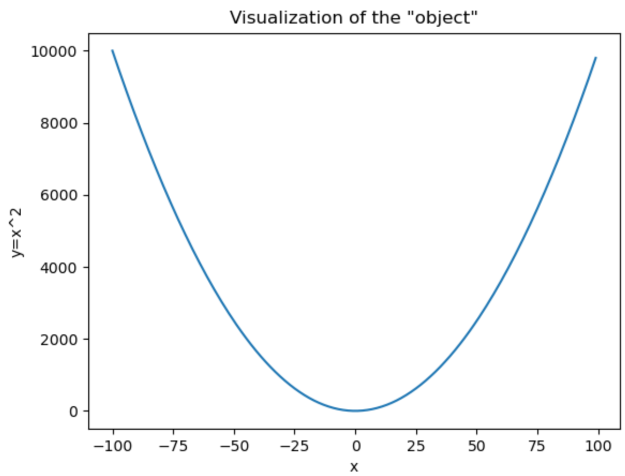
Coordinate Geometry

1.1 Introduction

Many of you have encountered some form of coordinate geometry in high school. For instance, the "standard" way to visualize a graph e.g. $f(x) = x^2$ is to visualize the points in 2-D space (x, y) where $y = x^2$. We give a demonstration in Python code.

```python
import matplotlib.pyplot as plt
import numpy as np
X=np.arange(-100,100) #create list of numbers from -100 to 100
Y= X**2 #calculate the square of at each x
plt.plot(X,Y) #plot all the pairs of points in 2d plane
plt.xlabel('x')
plt.ylabel('y=x^2')
plt.title('Visualization of the "object"')
plt.show()
plt.show()
```

$x \longrightarrow x^2$
the "graph" object



This is also known as the Cartesian plane, named after René Descartes who invented it in the 17th century.

1.2 Visualization of geometric objects

The Cartesian plane allows us to describe shapes with equations and perform calculations with them. We first define the playing field (the Cartesian plane and higher dimensional analogues) and the players.

Definition 1.1 (Real numbers)

The set of **real numbers**, denoted as $\mathbb{R}$, is (informally) the set of all the numbers that can be written out in decimal form.

Example 1.2

The following are real numbers:

1. The integers $0, \pm 1, \pm 2, \ldots$
2. Fractions in the form $\frac{a}{b}$, where a and $b \neq 0$ are integers.
3. Irrational numbers $\sqrt{2}$, π.

Remark. *The set of real numbers is known as a* ***complete field****. The definition of a complete field will be swept under the rug, but it guarantees a few things. The most important* *property: We will not "escape" the set by performing operations, possibly infinitely many.*

Definition 1.3 (N-dimensional space)

Let n be a positive integer. We denote the **n-dimensional real space** to be $\mathbb{R}^n$, consisting of all the n-tuples $(x_1, x_2, x_3, ..., x_n)$, where each x_j is a real number.
We call an n-tuple $(x_1, ..., x_n)$ a **point**, and two points $(x_1, ..., x_n)$ and $(y_1, ..., y_n)$ are equal if $x_j = y_j$ for all j-th entries of the tuples.

Remark. *We sometimes use* $\mathbf{x}$ *to denote* $(x_1, ..., x_n)$ *to make notation cleaner.*

1.2.1 Lines

Now that we have introduced the playing field of n-dimensional space, we can start translating the axioms of euclidean geometry to this coordinate system.

Definition 1.4 (Lines)

In euclidean geometry, a line is defined by two points. We let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. The **line** going from $\mathbf{x}$ to $\mathbf{y}$ is denoted $\overrightarrow{\mathbf{xy}}$.

What would this line look like? To get from $\mathbf{x}$ to $\mathbf{y}$, we have to traverse $y_1 - x_1$ units in the first coordinate, $y_2 - x_2$ units in the second, ..., $y_n - x_n$ in the last. We thus have a natural notation for the line $\overrightarrow{\mathbf{xy}}$.

$$\overrightarrow{\mathbf{xy}} = (y_1 - x_1, y_2 - x_2, ..., y_n - x_n).$$

This is very similar to a point as an n-tuple, but this is "spiritually" different to a point. This tuple represents the direction of line. One way to think of the correspondence between $(x_1, ..., x_n)$ point and $(x_1, ..., x_n)$ line is that $(x_1, ..., x_n)$ line is the line connecting $(0, 0, ...0)$ to $(x_1, ..., x_n)$ point. Because of this, we can identify a tuple as both the point and the line, and we call it a "vector" to abstract away from the actual geometric meaning.

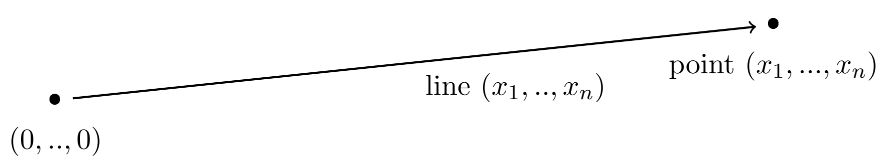


Figure 1.1: The correspondence between a line "vector" and a point "vector".

Notation. *Vectors are a very general notation of* $n-tuples$*. Depending on context, we use both of the following notations to denote the entries of* $\vec{v} \in \mathbb{R}^n$

- *"Ordered sets"* $(v_1, ..., v_n)$*, suitable dealing with points (and other geometric objects). It also looks cleaner when writing inline.*
- *"Column vectors"* $\begin{bmatrix} v_1 \\ . \\ . \\ . \\ v_n \end{bmatrix}$ *when the ordered sets are complicated to read, and when working with matrix algebra.*

1.2.2 Operation with lines

We need to translate a few more things from euclidean geometry.

Proposition 1.5

Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$, then

$$\overrightarrow{\mathbf{xy}} + \overrightarrow{\mathbf{yz}} = \overrightarrow{\mathbf{xz}},$$

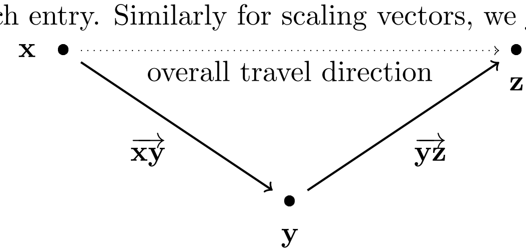
where $(a_1, ..., a_n) + (b_1, ..., b_n) \stackrel{\text{def}}{=} (a_1 + b_1, ..., a_n + b_n)$.

Proof.

$$\begin{aligned} \overrightarrow{\mathbf{xy}} + \overrightarrow{\mathbf{yz}} &= (y_1 - x_1, ..., y_1 - x_n) + (z_1 - y_1, ..., z_1 - y_n) \\ &= (y_1 - x_1 + z_1 - y_1, ..., y_n - x_n + z_n - y_n) \\ &= (z_1 - x_1, ..., z_n - x_n) \\ &= \overrightarrow{\mathbf{xz}} \end{aligned}$$

□

Geometrically, this means if you connect $\mathbf{x}$ to $\mathbf{y}$ to $\mathbf{z}$, the overall "direction of travel" you make is $\mathbf{x}$ to $\mathbf{z}$. This gives us a natural extension for addition of vectors by considering each entry. Similarly for scaling vectors, we just scale the entries along each dimension.



Definition 1.6 (Addition and scaling of vectors)

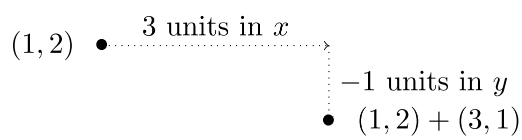
Let $\vec{a}, \vec{b}$ be two vectors in $\mathbb{R}^n$. We define the sum/difference of $\vec{a}$ and $\vec{b}$

$$\vec{a} + \vec{b} \stackrel{\text{def}}{=} (a_1 + b_1, ..., a_n + b_n), \quad \vec{a} - \vec{b} \stackrel{\text{def}}{=} (a_1 - b_1, ..., a_n - b_n)$$

and the scaling of $\vec{a}$ by a real number $c \in \mathbb{R}$

$$c\vec{a} \stackrel{\text{def}}{=} (ca_1, ca_2, ..., ca_n).$$

Remark. *Here we use the term "vectors", as we can in essence add points and lines together. How does one make sense of adding a line to a point? We can view this as translating the point along the path of the line, for instance, let us translate the point* $(1, 2)$ 3 *units in the first coordinate and* -1 *units in the second coordinate. This will give us* $(4, 1)$.



This way, we can write the line from $\vec{x}$ to $\vec{y}$ as $\vec{y} - \vec{x}$. The proof is a computational exercise.

Notation. *We now transferred for talking about points and the lines between points to addition. Therefore, we can overload the notation for points and lines as a vector* $\vec{v}$*, keeping in mind that they have the same arithmetic structure.*

In fact, most of our intuition for the real numbers translates to $\mathbb{R}$. For formality, we will list them here; in practice, we (almost always) take these properties for granted.

Proposition 1.7

Let $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$, and $a, b \in \mathbb{R}$. Then the following hold:

- *(Associativity)* $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$.
- *(Commutativity)* $\vec{u} + \vec{v} = \vec{v} + \vec{u}$.
- *(Identity)* The zero vector $\vec{0} \stackrel{\text{def}}{=} (0, 0, ..., 0) \in \mathbb{R}$ satisfies $\vec{v} + \vec{0} = \vec{v}$.
- *(Inverse)* The inverse of $\vec{v}$, $-\vec{v} \stackrel{\text{def}}{=} (-v_1, ..., -v_n)$ satisfies $\vec{v} + (-\vec{v}) = \vec{0}$.
- *(Scalar multiplication)* $a(b\vec{v}) = (ab)\vec{v}$.
- *(Scalar Identity)* $1\vec{v} = \vec{v}$.
- *(Distributivity 1)* $a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$.
- *(Distributivity 2)* $(a + b)\vec{v} = a\vec{v} + b\vec{v}$.

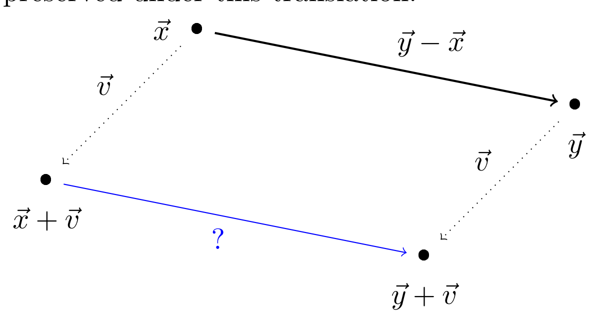
Remark. *All of these have good geometric intuition behind. For instance, the zero vector $\vec{0}$ is the "don't move" vector, corresponding to the point at the origin, or the "too short to be a line". The inverse of $\overrightarrow{\mathbf{xy}}$ is $\overrightarrow{\mathbf{yx}}$, where you go back from $\mathbf{x}$ to $\mathbf{y}$.*

Remark. *These 8 conditions are the axioms of a vector space. Later in the course, we will generalize the notion of vectors in $\mathbb{R}^n$ to other spaces (playing fields).*

Proposition 1.8

Lines are translation invariant. That is, for every $\vec{x}, \vec{y}, \vec{v} \in \mathbb{R}^n$, then the line from $\vec{x}$ to $\vec{y}$ is the same as the line from $\vec{x} + \vec{v}$ to $\vec{y} + \vec{v}$.

Let us illustrate what this statement is trying to convey. We have two points $\vec{x}$, $\vec{y}$; now we translate each of these points by $\vec{v}$, and we want the line between the points to be preserved under this translation.



The proof is one line: $(\vec{y} + \vec{v}) - (\vec{y} + \vec{v}) = \vec{y} - \vec{x} + \vec{v} - \vec{v} = \vec{y} - \vec{x}$. However, an immediate consequence of this is that we can "transport" vectors in space without distorting the vector. Colloquially, *5 miles South* to you describes the same direction and length as *5 miles South* to a person a few feet away. This justifies the way we visualize the correspondence between points and vectors - we "transport" the vectors to start from the origin $(0, ..., 0)$, and the end describes the point.

Remark. *Translation (and scaling) invariance is a property of Euclidean geometry. There are some exotic geometry systems that distort distance and direction through translation and scaling. One such example is the Poincaré metric.*

Another notion we can carry from Euclidean geometry is parallel lines. Here we not only define what it means for two vectors to be parallel (never touching), we also give a definition for two vectors to be parallel but point in opposite directions.

Definition 1.9 (Parallel and Antiparallel Vectors)

Let non-zero vectors $\vec{v}, \vec{w} \in \mathbb{R}^n$. We say that $\vec{v}$ and $\vec{w}$ are **parallel** if there is some $c > 0$ such that $\vec{v} = c\vec{w}$. We say that $\vec{v}$ and $\vec{w}$ are **antiparallel** if there is some $c < 0$ such that $\vec{v} = c\vec{w}$.

Example 1.10

Find an expression for the points on the (infinite) line passing through $P(1, 1, 0)$ and $Q(0, 2, 2)$.

Let M be a point on the line $\overrightarrow{PQ}$, then $\overrightarrow{PM}$ is parallel to $\overrightarrow{PQ}$ (or $M = P$). Either way, we would have some $c \in \mathbb{R}$ such that

$$\overrightarrow{PM} = c\overrightarrow{PQ}.$$

Indeed, we can confirm that any point represented as $P + c\overrightarrow{PQ}$ will be on the line, as $c\overrightarrow{PQ}$ is parallel to $\overrightarrow{PQ}$. One exception is to check for when $c = 0$, but we already know P is on the line. We thus have an expression for the line to be

$$(1, 1, 0) + c(-1, 1, 2)$$

for all $c \in \mathbb{R}$. Another way to put it, the line is the **image** of the function $f(c) = (1, 1, 0) + c(-1, 1, 2)$ on $\mathbb{R}$.

Remark. *This form of representing a line is known as a* ***parametric representation*** *or a* ***parametrization*** *of the line. We will give a more concrete definition of parametrization when we move beyond straight lines to "curvy" curves and surfaces in* **??**.

Remark. *Using the parametric representation for a line $l(t) = (P_1, P_2, ..., P_n) + t(v_1, v_2, ..., v_n)$, we can take slices/cross sections across each of the j-th coordinates, and get*

$$x_j = P_j + tv_j \implies t = \frac{x_j - P_j}{v_j}$$

is the equation of a line in 2d. P_j, v_j are fixed, viewing in the variables $y = x_j, x = t$, this is the equation of a line with y-intercept x_j and slope v_j. Setting the values of t for all slices to be equal,

$$t = \frac{x_1 - P_1}{v_1} = \frac{x_2 - P_2}{v_2} = ... = \frac{x_j - P_j}{v_j} = ... = \frac{x_n - P_n}{v_n}$$

*This form is known as the **symmetric equations** of a line.*

Example 1.11

Find an expression for the points on the line connecting **between** $P(1, 1, 0)$ to $Q(0, 2, 2)$.

This will be a segment of the line from the previous example, so the answer would be the same expression, but we limit the domain of f to be an interval on $\mathbb{R}$. Let us examine what f does to a few values of c.

c	very negative	$c = 0$	$c = 1/2$	$c = 1$	very positive
f(c)	very off the segment	P	midpoint of $\overline{PQ}$	Q	very off the line

As we move from very negative c to very positive c, we start very away from the line segment, reach P at $c = 0$, Q at $c = 1$, then move away from the line segment. Indeed the values $(1, 1, 0) + c(-1, 1, 2)$ will be on the line segment for $c \in [0, 1]$.

Finally, the nice algebraic properties for adding and scaling vectors gives us a natural way to understand these vectors as a sum of n component vectors, one for each dimension. These *special* vectors deserve a name.

Definition 1.12 (Standard Basis Vectors and Vector Decomposition)

We denote in $\mathbb{R}^n$, the **standard basis vectors**

$$\begin{aligned}\vec{e}_1 &\stackrel{\text{def}}{=} (1,0,0,...,0,0,0),\\ \vec{e}_2 &\stackrel{\text{def}}{=} (0,1,0,...,0,0,0),\\ \vec{e}_{n-1} &\stackrel{\text{def}}{=} (0,0,0,...,0,1,0),\\ \vec{e}_n &\stackrel{\text{def}}{=} (0,0,0,...,0,0,1),\end{aligned}$$

so that $\vec{v} = \sum_{i=1}^{n} v_i \vec{e}_i$. In $\mathbb{R}^3$, we sometimes write

$$\vec{i} \stackrel{\text{def}}{=} \vec{e}_1, \quad \vec{j} \stackrel{\text{def}}{=} \vec{e}_2, \quad \vec{k} \stackrel{\text{def}}{=} \vec{e}_3$$

to accommodate the physicists.

Exercises

1. Compute $5\vec{v} - 2\vec{w}$ and $-3\vec{w}$ for the following pairs of vectors:
   (a) $\vec{v} = 2\vec{i} + 3\vec{j}$, $\vec{w} = 4\vec{i} - 9\vec{j}$
   (b) $\vec{v} = (1, 2, -1)$, $\vec{w} = (2, -1, 0)$
   (c) $\vec{v} = -2\vec{e}_3 + 4\vec{e}_5$, $\vec{w} = \vec{e}_1 - 4\vec{e}_5$
   (d) $\vec{v} = (\cos t, \sin t)$, $\vec{w} = (\cos t)\vec{e}_2 - (\sin t)\vec{e}_1$
2. Find the vector $\overrightarrow{PQ}$ connecting $P(7, 2, 9)$ to $Q(-2, 1, 4)$. Put your answer in the form $a\vec{i} + b\vec{j} + c\vec{k}$. Give a parametric representation for the line $\overline{QP}$ in the form '$f(t) = \mathbf{a} + t\mathbf{b}$ for $t \in [c, d]$' this time the image moves from Q to P as the variable t increases.

1.3 Length, angles and projections

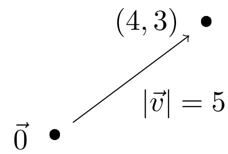
Definition 1.13 (Magnitude)

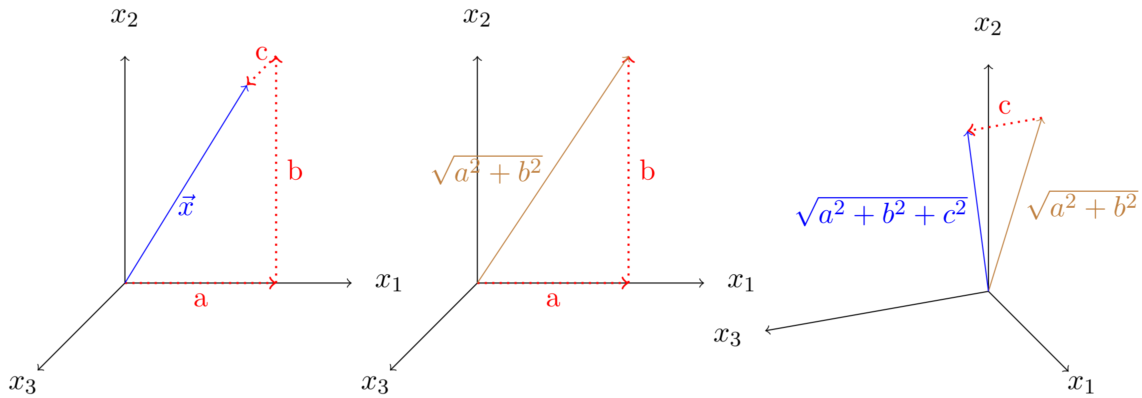
Let $\vec{v} \in \mathbb{R}^n$. The magnitude of $\vec{v}$ is denoted

$$|\vec{v}| \stackrel{\text{def}}{=} \sqrt{v_1^2 + v_2^2 + ... v_n^2}.$$

We build intuition through the lower dimensional cases. In $\mathbb{R}^2$, let us consider the point $(4, 3)$.

The magnitude of this vector is $\sqrt{3^2 + 4^2} = 5$. If this sounds very familiar, it is because this is indeed an application of Pythagorean theorem. In 3-dimensions, this still applies - take $\vec{x} = (a, b, c)$, we can traverse in each coordinate to apply Pythagorean theorem twice.





(a) $\vec{x} = (a, b, c)$. (b) Pyth. thm on a, b. (c) Reapply Pyth. thm.

Proposition 1.14

Let $\vec{v} \in \mathbb{R}^n, a \in \mathbb{R}$. Then
$$|a\vec{v}| = |a||\vec{v}|.$$

We have now transported the notions of length and scaling into the coordinate system, and this allows us to make "measurements" such as angles and area.

1.3.1 The Dot Product

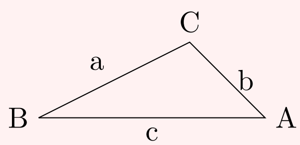
Example 1.15

Let points A=$(5, 8)$, B=$(-2, 7)$, O=$(4, 5)$. Find the angle $\angle AOB$.

To solve this using only information about the lengths, recall the Law of Cosines:

Theorem 1.16 (Law of Cosines)

For any triangle



,

$$c^2 = a^2 + b^2 - 2ab\cos(\angle BCA).$$

We now apply the Law of Cosines to $\triangle AOB$, so that

$$|\overrightarrow{AO}|^2 + |\overrightarrow{OB}|^2 - 2|\overrightarrow{AO}||\overrightarrow{OB}\ |\cos(\angle AOB) = |\overrightarrow{AB}|^2.$$

Plugging in the values, we solve

$$\begin{aligned}((4-5)^2+(5-8)^2)+((-2-4)^2+(7-5)^2)-2\sqrt{(4-5)^2+(5-8)^2}\\ \times\sqrt{(-2-4)^2+(7-5)^2}\cos(\angle AOB) = ((-2-5)^2+(7-8)^2).\end{aligned}$$

Simplifying, we get

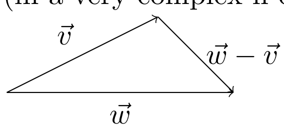
$$50 - 2\sqrt{10}\sqrt{40}\cos(\angle AOB) = 50 \implies \cos(\angle AOB) = 0.$$

So the angle is $\pi/2$.

Example 1.17

Find a closed form formula for the cosine of an angle formed by two vectors $\vec{v}, \vec{w} \in \mathbb{R}^n$.

Let the angle formed be θ. Using the intuition from 2-D space, we can form a triangle (in a very complex n-dimensional space). We write the Law of Cosine in terms of vectors



$$|\vec{v}|^2 + |\vec{w}|^2 - 2|\vec{v}||\vec{w}| = |\vec{w} - \vec{v}|^2.$$

This expands to

$$\sum_{i=1}^n v_i^2 + \sum_{j=1}^n w_j^2 - 2\sqrt{\sum_{i=1}^n v_i^2}\sqrt{\sum_{j=1}^n w_j^2}\ \cos(\theta) = \sum_{i=1}^n (w_i - v_i)^2.$$

Rearranging $(a-b)^2 = a^2 + b^2 - 2ab$,

$$2\sqrt{\sum_{i=1}^n v_i^2}\sqrt{\sum_{j=1}^n w_j^2}\ \cos(\theta) = \sum_{i=1}^n 2w_i v_i$$

so

$$\cos(\theta) = \frac{\sum_{i=1}^n w_i v_i}{\sqrt{\sum_{i=1}^n v_i^2}\sqrt{\sum_{j=1}^n w_j^2}}.$$

Definition 1.18 (Dot Product)

Let $\vec{v}, \vec{w} \in \mathbb{R}^n$, the **dot product** between $\vec{v}$ and $\vec{w}$ is

$$\vec{v} \cdot \vec{w} \overset{\text{def}}{=} \sum_{i=1}^n v_i w_i.$$

Proposition 1.19

Let $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$, $c \in \mathbb{R}$. The dot product satisfies the following properties:

- *(Symmetry)* $\vec{v} \cdot \vec{w} = \vec{w} \cdot \vec{v}$.
- *(Linearity 1)* $(c\vec{v}) \cdot \vec{w} = c(\vec{v} \cdot \vec{w})(= \vec{v} \cdot (c\vec{w})$ by symmetry).
- *(Linearity 2)* $(\vec{v} + \vec{u}) \cdot \vec{w} = \vec{v} \cdot \vec{w} + \vec{u} \cdot \vec{w}$.
- *(Positive definiteness)* $\vec{v} \cdot \vec{v} \geq 0$, with equality if and only if $\vec{v} = \vec{0}$.

Proof. **TODO: write the proof, or put it as an exercise.** □

Corollary 1.20:

1. $|\vec{v} - \vec{w}|^2 = |\vec{v}|^2 + |\vec{w}|^2 - 2\vec{v} \cdot \vec{w}$.
2. If $\vec{v}, \vec{w} \neq \vec{0}$, the angle between $\vec{v}$ and $\vec{w}$ is

   $$\cos^{-1}\left(\frac{\vec{v} \cdot \vec{w}}{|\vec{v}||\vec{w}|}\right).$$

   In particular, the angle is $\pi/2$ when $\vec{v} \cdot \vec{w} = 0$.
3. $|\vec{v} \cdot \vec{w}| \leq |\vec{v}||\vec{w}|$.

Since we can make sense of "angles" in higher dimensions, it is natural to generalize the notion of two vectors "perpendicular" to each other.

Definition 1.21 (Orthogonality)

Let $\vec{v}, \vec{w} \in \mathbb{R}^n$. We say $\vec{v}$ and $\vec{w}$ are **orthogonal** to each other if $\vec{v} \cdot \vec{w} = 0$.

Example 1.22

Find all vectors $\vec{v} \in \mathbb{R}^3$ such that $\vec{v} \cdot (-3\vec{i} + 2\vec{j} + 6\vec{k}) = 49$. Sketch the locus of corresponding points in $\mathbb{R}^3$.

The condition expands to

$$\begin{aligned} -3v_1 + 2v_2 + 6v_3 &= 49 \\ \implies v_3 &= \frac{49 + 3v_1 - 2v_2}{6} \end{aligned}$$

so that the set of all vectors is

$$\left\{ s\vec{i} + t\vec{j} + \left(\frac{49 + 3s - 2t}{6} \right) \vec{k} | s, t \in \mathbb{R} \right\}.$$

To plot this in 3D space, notice that this describes the equation of a plane $z = (49 + 3x - 2y)/6$. We pick any three points (of course we want those that are easy to calculate) $(0, 0, 49/6)$, $(0, 49/12, 0)$, $(-49/18, 0, 0)$.

```
def function_to_plot(X,Y):
return (49+3*X-2*Y)/6
#create figure
fig=plt.figure(figsize=(10,8))
ax = fig.add_subplot(projection='3d')
ax.view_init(elev=10, azim=120) #change view 1
#ax.view_init(elev=10, azim=0) #change view 2
X,Y=np.meshgrid(np.linspace(-10,10,10), np.linspace(-10,10,10))
Z=function_to_plot(X,Y)
surface=ax.plot_surface(X, Y, Z, alpha=0.5,
label='graph surface')
ax.set_xlabel('x')
ax.set_ylabel('y')
ax.set_zlabel('z')
#plot the curve and the cross section to integrate
ax.scatter(0,0,0,color='red',label='origin')
ax.text(0,0,0,'origin')
ax.scatter(-3,2,6,color='black')
ax.text(-3,2,6,'(-3,2,6)')
plt.legend()
plt.show()
```

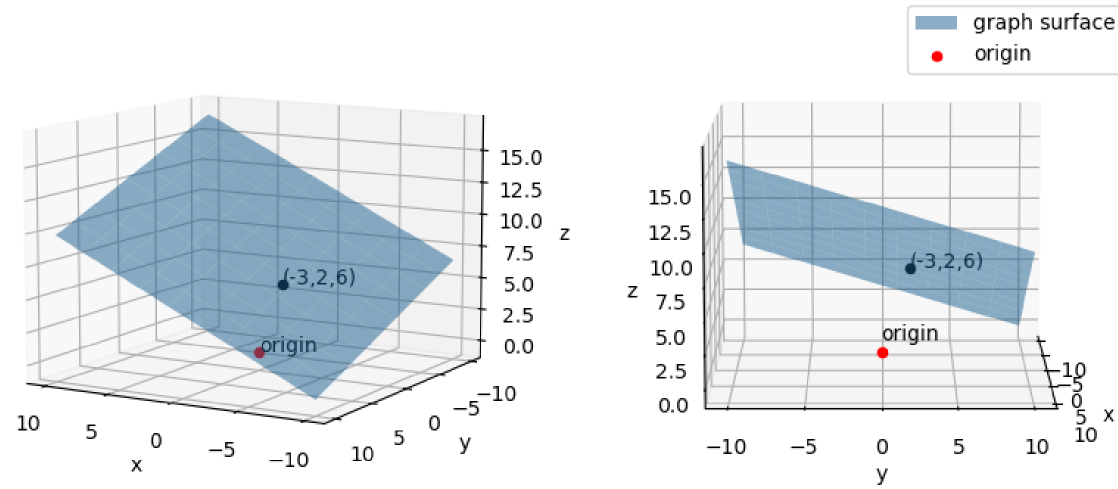


Figure 1.3: Two views of the plane

Interestingly, we see that $(-3, 2, 6)$ - the point corresponding to our vector - is a point on the plane!

Remark. *The form*

$$\left\{ s\vec{i} + t\vec{j} + \left(\frac{49 + 3s - 2t}{6} \right) \vec{k} | s, t \in \mathbb{R} \right\}$$

is a ***parametrization*** *of the plane described by* $-3x + 2y + 6z = 49$*, viewing this as a function* $\mathbb{R}^2 \to \mathbb{R}^3$.

1.3.2 Projection

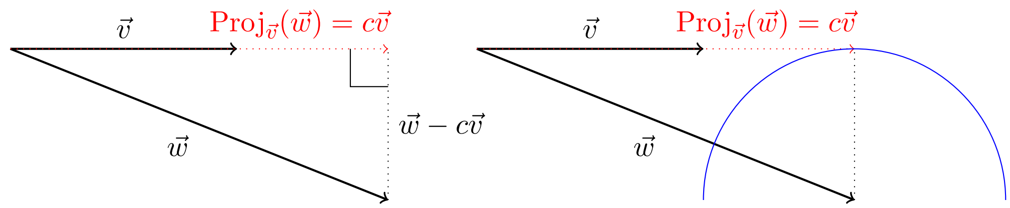
Definition 1.23 (Projection)

Let $\vec{v}, \vec{w} \in \mathbb{R}^n$, $\vec{v} \neq \vec{0}$, and $c \in \mathbb{R}$, such that $\vec{v} \cdot (\vec{w} - c\vec{v}) = 0$ (or equivalently, $\vec{w} - c\vec{v}$ orthogonal to $\vec{v}$).
We say that $c\vec{v} \overset{\text{def}}{=} \text{Proj}_{\vec{v}}(\vec{w})$ is the **vector projection of $\vec{w}$ on $\vec{v}$**.

To build intuition, it is always helpful to start with lower dimensions. We take a plane through $\vec{v}$ and $\vec{w}$. Looking at this slice, we can work in 2D.

We see that $c\vec{v}$ is the the point on the extension of $\vec{v}$ such that it is closest to the point $\vec{w}$. To geometrically show this idea, we draw a circle centered at $\vec{w}$ with radius $|\vec{w} - c\vec{v}|$.

The line generated by $\vec{v}$ is tangent to this circle, so the contact point at $c\vec{v}$ is indeed the closest point to $\vec{w}$.



We also see from this geometric construction that 1) the projection $c\vec{v}$ is unique thus well defined, 2) $c|\vec{v}| = |\vec{w}|\cos\theta$

Proposition 1.24

The projection is unique and is given by

$$\text{Proj}_{\vec{v}}(\vec{w}) = \frac{\cos\theta|\vec{w}|}{|\vec{v}|} = \frac{\vec{v}\cdot\vec{w}}{|\vec{v}|^2} = (\frac{\vec{v}\cdot\vec{w}}{\vec{v}\cdot\vec{v}})\vec{v}.$$

This is the first time we show that something is unique. The standard argument goes as follows: Suppose another object a' satisfies all the properties you want for a. Then you can show $a = a'$, meaning every object that satisfies the properties is a, or equivalently, a is unique.

So what happens if we pick two points P, P' on $\vec{v}$ such that both make a right angle when connected to $\vec{w}$? We would have constructed a triangle $\triangle PP'\vec{w}$ with two right angles!

Proof. Let $c, c' \in \mathbb{R}$ such that $\vec{p}_1 = c\vec{v}$ and $\vec{p}_2 = c'\vec{v}$ are both satisfy the definition of $\text{Proj}_{\vec{v}}(\vec{w})$. We want to show that $c = c'$, so we consider $\vec{u} = (c - c')\vec{v}$, the line between the two projections.
By a corollary in 1.3.1, we have two equations

$$|\vec{w} - \vec{p}_1|^2 = |\vec{w} - \vec{p}_2|^2 + |\vec{u}|^2$$
$$|\vec{w} - \vec{p}_2|^2 = |\vec{w} - \vec{p}_1|^2 + |\vec{u}|^2$$

We can solve for $|\vec{u}| = 0$, and by positive definiteness, $\vec{u} = \vec{0}$ and $c - c' = |\vec{u}|/|\vec{v}| = 0$.
To show our formula for projection works, we can just compute that

$$\vec{v}\cdot\left(\vec{w} - \left(\frac{\vec{v}\cdot\vec{w}}{\vec{v}\cdot\vec{v}}\right)\vec{v}\right) = \vec{v}\cdot\vec{w} - \left(\frac{\vec{v}\cdot\vec{w}}{\vec{v}\cdot\vec{v}}\right)(\vec{v}\cdot\vec{v}) = 0.$$

□

Exercises

3. Determine if the following pairs of vectors are orthogonal:

   (a) $3\vec{e}_1 + 4\vec{e}_2$, $-4\vec{e}_1 + 3\vec{e}_2$

   (b) $(4, -1, 2)$, $(3, 0, -6)$

   (c) $3\vec{i} - 2\vec{j}$, $-2\vec{i} - 4\vec{k}$

   (d) $\vec{e}_1 + 3\vec{e}_3 + 5\vec{e}_5 + ... + (2k-1)\vec{e}_{2k-1}$, $2\vec{e}_2 + 4\vec{e}_4 + ... + (2k)\vec{e}_{2k}$

4. Refer to the plane in the example in the previous section 1.3.1. Show that $\text{Proj}_{-3\vec{i}+2\vec{j}+6\vec{k}}(\vec{v})$ is the same for any vector $\vec{v}$ such that $\vec{v} \cdot (-3\vec{i} + 2\vec{j} + 6\vec{k}) = 49$, and compute this projection.

5. *(Geometry of a methane molecule)* Place four points $P(0,0,0)$, $Q(1,1,0)$, $R(1,0,1)$, $S(0,1,1)$ in $\mathbb{R}^3$.

   (a) Compute the distance between any two points of $PQRS$ and show that all 6 pairs are the same. This means that $PQRS$ forms a regular tetrahedron.

   (b) Verify that the geometric center of the tetrahedron $O(1/2, 1/2, 1/2)$ is equidistant to all of the vertices of the tetrahedron.

   (c) Compute the angle between two edges of the tetrahedron, rounded to 2 decimal places. By symmetry, you only have to compute one angle.

   (d) Compute the angle $\angle POQ$, rounded to 2 decimal places. Does this angle remind you of something from Chemistry?

1.4 Cross Product

Definition 1.25 (Cross Product)

Let $\vec{v}, \vec{w} \in \mathbb{R}^3$, we define the **cross product** of $\vec{v}$ and $\vec{w}$ to be

$$\vec{v} \times \vec{w} \stackrel{\text{def}}{=} \begin{bmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{bmatrix}.$$

Remark. *Later on, we will introduce determinants of a matrix. The cross product can be* *understood as the determinant of the 'matrix'*

$$\begin{bmatrix} \vec{i} & v_1 & w_1 \\ \vec{j} & v_2 & w_2 \\ \vec{k} & v_3 & w_3 \end{bmatrix}.$$

This definition seems a bit unmotivating, so let us work through some examples.

Example 1.26

Compute the 9 cross products for each pair of the standard basis vectors in $\mathbb{R}^3$.

With some (heavy) computation, we find

$$\begin{array}{lll} \vec{i} \times \vec{i} = \vec{0} & \vec{i} \times \vec{j} = \vec{k} & \vec{i} \times \vec{k} = -\vec{j} \\ \vec{j} \times \vec{i} = -\vec{k} & \vec{j} \times \vec{j} = \vec{0} & \vec{j} \times \vec{k} = \vec{i} \\ \vec{k} \times \vec{i} = \vec{j} & \vec{k} \times \vec{j} = -\vec{i} & \vec{k} \times \vec{k} = \vec{0} \end{array}$$

Importantly, the cross product is **antisymmetric**, meaning $\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}$. You can see special cases with the standard basis from above and confirm the general case in an exercise.

Example 1.27

Compute $\vec{v} \cdot (\vec{v} \times \vec{w})$ and $\vec{w} \cdot (\vec{v} \times \vec{w})$.

Again with some heavy computation,

$$
\begin{aligned}
\vec{v}\cdot(\vec{v}\times\vec{w}) &= \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \cdot \begin{bmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{bmatrix} \\
&= v_1v_2w_3 - v_1v_3w_2 + v_2v_3w_1 - v_2v_1w_3 + v_3v_1w_2 - v_3v_2w_1 \\
&= 0, \\
\vec{w}\cdot(\vec{v}\times\vec{w}) &= \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \cdot \begin{bmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{bmatrix} \\
&= w_1v_2w_3 - w_1v_3w_2 + w_2v_3w_1 - w_2v_1w_3 + w_3v_1w_2 - w_3v_2w_1 \\
&= 0.
\end{aligned}
$$

Which means $\vec{v}\times\vec{w}$ is orthogonal to $\vec{v}$ and $\vec{w}$!

Example 1.28

Compute $|\vec{v}\times\vec{w}|^2 + (\vec{v}\cdot\vec{w})^2$.

We have

$$
\begin{aligned}
|\vec{v}\times\vec{w}|^2 + (\vec{v}\cdot\vec{w})^2 =& (v_2w_3 - v_3w_2)^2 + (v_3w_1 - v_1w_3)^2 + (v_1w_2 - v_2w_1)^2 \\
& + (v_1w_1 + v_2w_2 + v_3w_3)^2 \\
=& (v_1w_1)^2 + (v_1w_2)^2 + (v_1w_3)^2 \\
& + (v_2w_1)^2 + (v_2w_2)^2 + (v_2w_3)^2 \\
& + (v_3w_1)^2 + (v_3w_2)^2 + (v_3w_3)^2 \\
=& (v_1^2 + v_2^2 + v_3^2)(w_1^2 + w_2^2 + w_3^3) \\
=& |\vec{v}|^2|\vec{w}|^2.
\end{aligned}
$$

Now we substitute $\vec{v}\cdot\vec{w} = |\vec{v}||\vec{w}|\cos\theta$, we get

$$|\vec{v}\times\vec{w}| = |\vec{v}||\vec{w}|\sqrt{1-\cos^2\theta} = |\vec{v}||\vec{w}|\sin\theta.$$

Where we know $\sin\theta \geq 0$ as θ is between 0 and π.

The value $|\vec{v}||\vec{w}|\sin\theta$ has a nice geometric meaning. It is the area spanned by the vectors $\vec{v}$ and $\vec{w}$.

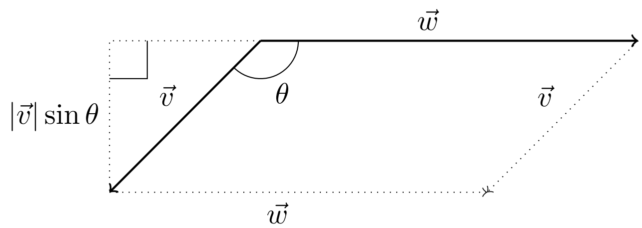


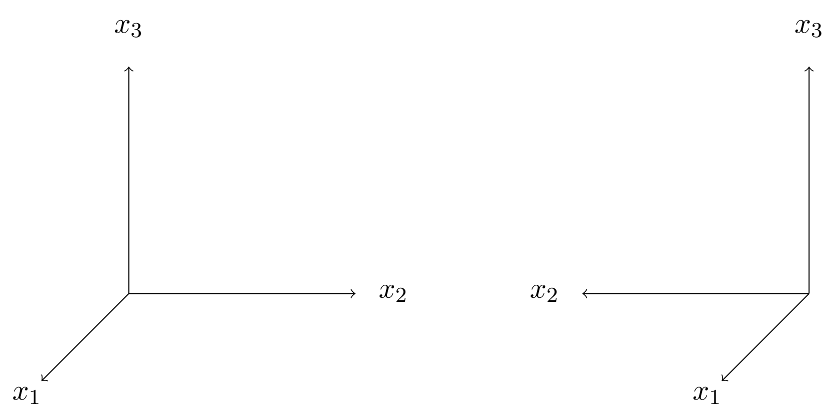
Figure 1.5: The area of a parallelogram formed by these vectors is the magnitude of the vector $\vec{v} \times |\vec{w}|$

Proposition 1.29

The direction of $\vec{v} \times \vec{w}$ is determined by the **right-hand rule** as follows:

> Using the right hand, align the index finger with the direction $\vec{v}$, and the middle finger with the direction of $\vec{w}$. Extend the thumb so that it is perpendicular to both the index finger and the middle finger. The thumb is pointing in the direction of $\vec{v} \times \vec{w}$.

This is a byproduct of the convention we use. In $\mathbb{R}^3$ we use what is known as a right-handed coordinate system - the vectors $\vec{i}$, $\vec{j}$, and $\vec{k}$ align with the first three fingers of the right hand respectively. If we used a left-handed coordinate system, the rule would be left-handed instead. We now have a few properties about the cross product from our



(a) Right handed coordinate system (b) Left handed coordinate system

computation, the first you will verify on your own:

Proposition 1.30

Let $\vec{v}, \vec{w}, \vec{u} \in \mathbb{R}^n$, then

- latex sucks. I hate using tikz.
- *(distributivity)* $\vec{v} \times (\vec{u} + \vec{w}) = \vec{v} \times \vec{u} + \vec{v} \times \vec{w}$ and $(\vec{v} + \vec{u}) \times \vec{w} = \vec{v} \times \vec{w} + \vec{v} \times \vec{u}$.
- *(anti-symmetry)* $\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}$.
- $\vec{v} \times \vec{w}$ is orthogonal (perpendicular in 3D space) to $\vec{v}$ and $\vec{w}$.
- The magnitude $|\vec{v} \times \vec{w}|$ is given by $|\vec{v}||\vec{w}| \sin\theta$, with θ being the angle between $\vec{v}$ and $\vec{w}$, so
  (i) The magnitude $|\vec{v} \times \vec{w}|$ also corresponds to the area of the parallelogram formed by $\vec{v}$ and $\vec{w}$.
  (ii) If $\vec{v}$ and $\vec{w}$ are parallel or antiparallel, $\vec{v} \times \vec{w} = \vec{0}$.

Remark. *Using distributivity of the cross product, you only need to memorize the cross product of the basis vectors, and write* $\vec{v} \times \vec{w} = \sum_{i=1}^{3} \sum_{j=1}^{3} v_i w_i (\vec{e}_i \times \vec{e}_j)$.

1.4.1 Triple products

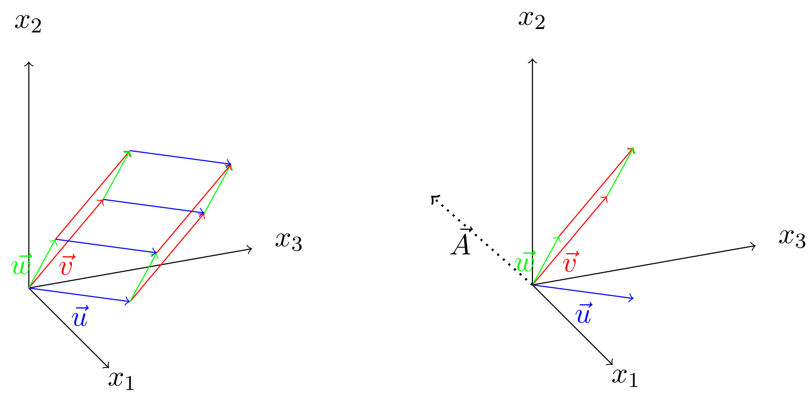
Example 1.31

Find the volume of the parallelepiped formed from $\vec{v} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, $\vec{w} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\vec{u} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

The parallelepiped is a generalization of a parallelogram to higher dimensions. Using combinations of $\vec{v}, \vec{w}, \vec{u}$, you can make the frame of a 3d solid. As in the figure on the left, edges of the same color correspond to the same vector (and thus parallel). The volume of this solid is still *base* $\times$ *height*, where the base is a 2D parallelogram formed by two vectors and the height is determined by third vector. We make an arbitrary decision and set the base to be $\vec{v}$ and $\vec{w}$. (setting any two vectors would give the same result in the end!) The area of this base is given by the cross product

$$\vec{A} = \vec{v} \times \vec{w} = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}.$$

Now we can determine the height of the paralellepiped from $\vec{u}$. We want to isolate the component of $\vec{v}$ that is orthogonal to the base. Equivalently, we want to find the component of $\vec{u}$ that is pointing in the direction of $\vec{A}$, a vector that is orthogonal to both $\vec{v}$ and $\vec{u}$!



(a) The parallelepiped, draw in an optical illusion fashion.

(b) We want to get the projection of $\vec{u}$ on $\vec{A}$.

The volume is thus

$$|\mathrm{Proj}_{\vec{A}}(\vec{u})||\vec{v}| = |\frac{\vec{u}\cdot\vec{A}}{|\vec{A}|^2}| \times |\vec{A}| \times |\vec{A}| = |\vec{u}\cdot\vec{A}| = 2.$$

Proposition 1.32

The volume of the parallelpiped formed from $\vec{v}, \vec{w}, \vec{u}$ is

$$|(\vec{v}\times\vec{w})\cdot\vec{u}|$$

Remark. *This is also the expression of the (absolute value of) determinant of*

$$\begin{bmatrix} \vec{v} & \vec{w} & \vec{u} \end{bmatrix}$$

where $\vec{v}, \vec{w}, \vec{u}$ are written as column vectors. Using properties of the determinant (later chapters), you can show cycling the three vectors does not change the volume. (i.e. you can calculate using what order of the three vectors you want)

Remark. *You may notice that the expression$(\vec{v} \times \vec{w}) \cdot \vec{u}$ can take on negative volumes. In this case, the three vectors (taken in order) do not follow the right-hand rule. For instance, in the last example, $\vec{u}$ points in the 'opposite' direction as $\vec{A}$.*

Exercises

6. Find $\vec{v} \times \vec{w}$ for the following:
   (a) $\vec{v} = (4, -2, 0)$, $\vec{w} = (2, 1, -1)$
   (b) $\vec{v} = (3, 3, 3)$, $\vec{w} = (4, -3, 2)$
   (c) $\vec{v} = 2\vec{i} + 3\vec{j} + 4\vec{k}$, $\vec{w} = \vec{i} - 3\vec{j} + 4\vec{k}$
7. Find the areas for the following shapes:
   (a) The parallelogram with vertices $P(0, 0, 0), Q(1, 1, 0), R(1, 2, 1), S(0, 1, 1)$.
   (b) The triangle with vertices $A(1, 9, 3), B(-2, 3, 0), C(3, -5, 3)$.
8. Find the volume of the parallelepiped formed from the vectors $\vec{v} = (1, 1, 0), \vec{w}(0, 2, -2), \vec{u} = (1, 0, 3)$.
9. A triangular kite has vertices $P(0, 0, 10), Q(2, 1, 10), S(0, 3, 12)$ and is displaced by the wind at a velocity of $(20\vec{i} + 6\vec{j} + 4\vec{k})/s$
   (a) Find the area of the kite.
   (b) After 1/2 seconds, find the volume of the space swept by the kite. (leave the answers in $[units]^3$)

1.5 Applications - Geometry of lines and planes

Definition 1.33 (Relations between lines)

For two (infinitely extending) lines in $\mathbb{R}^n$ parametrized in s and t respectively as $l_1 = P + t\vec{v}$, $l_2 = Q + s\vec{w}$, we say the lines are

- **Parallel**, if the $\vec{v}$ and $\vec{w}$ are parallel or antiparallel.
- **Intersecting**, if l_1 and l_2 exactly one point on both l_1 and l_2.
- **Skew**, if l_1 and l_2 are not parallel/antiparallel or intersecting.

Remark. *Lines do not have direction, so there usually is no need to distinguish between parallel and antiparallel lines. One may extend the definition of antiparallel to lines from $\vec{v}$ and $\vec{w}$.*

Proposition 1.34

Determination of parallel lines are independent of parametrization. Concretely,

> Let $P_1 + t_1\vec{v}_1$ and $P_2 + t_2\vec{v}_2$ be two parametrizations of l_1, $Q_1 + s_1\vec{w}_1$, $Q_2 + s_2\vec{w}_2$ be two parametrizations of l_2. If $\vec{v}_1 = c_1\vec{w}_1$ for some $c_1 \in \mathbb{R}$, then $\vec{v}_2 = c_2\vec{w}_2$ for some (possibly different)$c_2 \in \mathbb{R}$.

The proof is not very enlightening. However, the result of this guarantees that our definition of parallel lines is precise.

Proof. The idea is to show that $\vec{v}_1$ is a scalar multiple of $\vec{v}_2$, and by the same logic $\vec{w}_1$ is a scalar multiple of $\vec{w}_2$. Since all vectors are non-zero in the parametrization, we will get the result of $\vec{v}_2$ a scalar multiple of $\vec{w}_2$.

To show $\vec{v}_1 = k\vec{v}_2$ for some k, we can pick two distinct points A, B on l_1. From the parametriation we can get $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$ such that $P_1 + \alpha_1\vec{v}_1 = A = P_2 + \alpha_2\vec{v}_2$ and $P_1 + \beta_1\vec{v}_1 = B = P_2 + \beta_2\vec{v}_2$. Therefore we get the vector

$$\begin{aligned}\overrightarrow{AB} &= (\beta_1 - \alpha_1)\vec{v}_1,\\ \overrightarrow{AB} &= (\beta_2 - \alpha_2)\vec{v}_2.\end{aligned}$$

As we picked distinct points A and B, we can conclude $\overrightarrow{AB} \neq \vec{0}$ and thus $\beta_1 - \alpha_1 \neq 0, \beta_2 - \alpha_2 \neq 0$, so that

$$\vec{v}_1 = \frac{\beta_2 - \alpha_2}{\beta_1 - \alpha_1}\vec{v}_2.$$

□

Example 1.35

Determine whether the lines parametrized by $l_1(t) = (1, 2, 1) + t(1, 3, -2)$ and $l_2(t) = (3, 1, 0) + t(-2, -6, 4)$ are parallel, intersecting, or skew. Confirm that l_1 and l_2 describe two different lines.

We notice that $-2 \times (1, 3, -2) = (-2, -6, 4)$, so these lines are parallel. To confirm that these two lines are not the same, we notice that $(1, 2, 1)$ is a point on l_1, but if we attempt to solve

$$(1, 2, 1) = l_2(t) = (3, 1, 0) + t(-2, -6, -4) \implies (-2, 1, 1) = t(-2, -6, -4)$$

which no t can solve! Specifically, the first coordinate forces $t = 1$ and the second coordinate forces $t = -6$.

Example 1.36

Determine whether the lines parametrized by $l_1(t) = (1,2,1) + t(1,3,-2)$ and $l_2(t) = (0,3,9) + t(0,2,3)$ are parallel, intersecting, or skew.

$(1,3,2)$ is not a multiple of $(0,2,3)$, so the lines are not parallel. We might be tempted to solve for $l_1(t) = l_2(t)$ to check for intersection, but this misses a lot of cases! We need to compare all the points of l_1 with all the points of l_2, so we need two independent variables to describe where we are on each of the lines. That is, we solve for s, t in $l_1(t) = l_2(s)$,

$$\begin{aligned} (1,2,1) + t(1,3,-2) &= (0,3,9) + s(0,2,3) \\ \implies (t+1, 3t+2, -2t+1) &= (0, 2s+3, 3s+9) \\ \implies t = -1 \text{ and } 3t+2 &= 2s+3 \text{ and } -2t+1 = 3s+9 \end{aligned}$$

$t = -1, s = -2$ solves this system of equations. We can plug in t and s in our original parametrization to find $(0,1,3)$ is indeed a point on both l_1 and l_2. We would have missed this if we set $l_1(t) = l_2(t)$!

Example 1.37

Determine whether the lines parametrized by $l_1(t) = (1,2,1) + t(1,3,-2)$ and $l_2(t) = (0,3,8) + t(0,2,3)$ are parallel, intersecting, or skew.

We repeat the same process as above to see that the lines are not parallel and solve for

$$\begin{aligned} (1,2,1) + t(1,3,-2) &= (0,3,8) + s(0,2,3) \\ \implies (t+1, 3t+2, -2t+1) &= (0, 2s+3, 3s+8) \\ \implies t = -1 \text{ and } 3t+2 &= 2s+3 \text{ and } -2t+1 = 3s+8 \end{aligned}$$

This time, we do not have a solution - the first two equations forces $t = -1, s = -2$, and this does not solve the third. We therefore do not have a point of intersection, and the lines are skew.

Example 1.38

Determine if $P(5,6,9), Q(7,9,15), R(13,18,33)$ are **colinear** i.e. if they lie on the same line.

With the machinery we have built up, there are multiple ways to check if P, Q, R form a straight line. Here are a few ideas:

1. Check that $\overrightarrow{PQ}$ is parallel/antiparallel to $\overrightarrow{PR}$. Because the lines defined by $\overline{PQ}$ and $\overline{PR}$ are parallel and share a same point P, they are the same line.

2. Use the dot product to calculate the angle $\angle PQR = \pi$.

3. Use the cross product to calculate that $\overrightarrow{PQ} \times \overrightarrow{PR} = \vec{0}$. This means the triangle with vertices P, Q, R has no area and thus is a degenerate triangle.

Definition 1.39 (Characterization of Planes)

In euclidean geometry, planes can be characterized by any of the following ways:

- For **any three non-colinear points** P_1, P_2, P_3, there is a unique plane passing through P_1, P_2, P_3.
- For **any pair of intersecting lines**, there is a unique plane that contains both.
- For **a line** l **and a point** P, there is a unique plane that contains P and is perpendicular to l.
- For **a line** l **and a point** P **not on** l, there is a unique plane that contains both l and P.

Remark. *The third characterization is the hardest to visualize at first, but is also the easiest to describe with the analytical tools we have built towards. We can refer to example 1.22 in 1.3.1. The plane sketched is the unique plane that contains* $(-3, 2, 6)$*, such that each vector in the plane is orthogonal to* $(-3, 2, 6)$*, so the line parametrized by* $l(t) = t(-3, 2, 6)$ *is perpendicular to the plane at the point of intersection* $(-3, 2, 6)$*.*

Example 1.40

In $\mathbb{R}^3$, determine the equation of the plane that contains $P_0(x_0, y_0, z_0)$ and is perpendicular to the line parametrized by $l(t) = Q_0 + t\vec{N}$, $\vec{N} = (a, b, c)$.

First we can exploit **translation invariance** of $\mathbb{R}^3$ and move the line to $\tilde{l} = P_0 + t\vec{N}$. Because l and $\tilde{l}$ are parallel, any line perpendicular to l will also be perpendicular to $\tilde{l}$. Then by the characterization of a plane, any P on the plane satisfies the orthogonality relation $\overrightarrow{PP_0} \cdot \vec{N} = 0$. Expanding this, we get the equation

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \implies ax + by + cz = ax_0 + by_0 + cz_0.$$

Definition 1.41 (Normal form of a plane)

Let $a, b, c, d \in \mathbb{R}$, with at least one of $a, b, c \neq 0$. The equation of a plane in ^{3} written as

$$ax + by + cz = d$$

is called a **normal form** of the equation of a plane.

Remark. *As all planes can be characterized this method, all equations of planes can be put in normal form.*

Remark. *The normal form of a plane is not unique. Pick your favorite non-zero number α, the equation $\alpha ax + \alpha by + \alpha c_z = \alpha d$ describes the same plane.*

Theorem 1.42 (Normal vectors of planes)

When written in normal form, the plane is perpendicular to the vector $\vec{N} = (a, b, c)$. We call this vector $\vec{N}$ the **normal vector**.

A problem in the exercise will guide you through the proof of this. The intuition behind the proof is the reverse direction of the equation $\overrightarrow{PP_0} \cdot \vec{N} = 0$ we derived from the last example.
We will now apply this theorem in a few examples.

Example 1.43

Find the equation of the plane that passes through the points $P(1, 0, 0), Q(0, 1, 0), R(0, 0, 1)$.

Method 1: One sees that (by coincidence) the sum of coordinates of each point are equal to 1, so immediately writes down $x + y + z = 1$. Despite being the fastest method, this is somewhat inconsistent.

Method 2: We set $ax + by + cz = d$ to be the equation, and plug in values for P, Q, R, giving the system of equations

$$\begin{aligned} a + 0b + 0c &= d \\ 0a + b + 0c &= d \\ 0a + 0b + c &= d \end{aligned}$$

This is an **underdetermined system**, meaning there are fewer equations than unknowns. The best we can say is that $a = b = c = d$. However, setting $a = b = c = d = k$ works for all $k \neq 0$, further confirming that the normal form is not unique. This method is reasonably fast when the system of equations are simple. When there are more non-zero coefficients, solving the system takes more time.

Method 3: We can compute the direction of the normal vector $\vec{N}$ of this plane. By the characterization of planes, $\vec{N}$ is orthogonal to $\overrightarrow{PQ}$ and $\overrightarrow{PR}$, two vectors in the plane. We can thus write $\vec{N}$ as a multiple of the cross product

$$\overrightarrow{PQ} \times \overrightarrow{PR} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \times \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

. Using this normal vector (or any multiple of it) and applying the theorem, we get $x + y + z = d$, and substituting P into the equation will give $x + y + z = 1$. This method is more general and consistent, as the number of operations in a cross product is constant.

Example 1.44

On the plane given by $ax + by + cz = d$, find the point on the plane that is closest to the origin.

Let r be the distance, we draw a sphere centered at the origin with radius r. This sphere is thus tangent to the surface at one point, and the vector corresponding to this point will be perpendicular to the plane. By the theorem, we can denote the point $P(\alpha a, \alpha b, \alpha c)$, with α a constant to be determined. Substituting P into the equation,

$$\alpha(a^2 + b^2 + c^2) = d \implies \alpha = \frac{d}{a^2 + b^2 + c^2}$$

The closest distance from the origin is

$$|(\alpha a, \alpha b, \alpha c)| = |\alpha|\sqrt{a^2 + b^2 + c^2} = \left|\frac{d}{\sqrt{a^2 + b^2 + c^2}}\right|$$

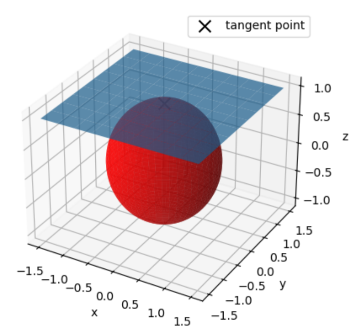


Figure 1.8: A toy example with the unit sphere of radius 1 and the plane described by $z = 1$.

at the point

$$\left(\frac{da}{\sqrt{(a^2 + b^2 + c^2)}}, \frac{db}{\sqrt{(a^2 + b^2 + c^2)}}, \frac{dc}{\sqrt{(a^2 + b^2 + c^2)}}\right)$$

1.5.1 Intersection of planes

Example 1.45

Find the intersection of the planes given by

$$6x + 2y - z = 2$$
$$\text{and } x - 2y + 3z = 5.$$

Method 1: We can solve the system of equations to find the line of intersection, and arrive at the set of symmetric equations. A similar method would be to solve for two points on the line of intersection, then get the parametric equation. **TODO, only if i feel like it**

Method 2: The line lies on the first plane, so is perpendicular to its normal vector $(6, -2, -1)$. The line also lies on the second plane, so is perpendicular to its normal vector $(1, -2, 3)$. Using the cross product, the line should point in the direction of

$$\begin{bmatrix} 6 \\ 2 \\ -1 \end{bmatrix} \times \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ -19 \\ -14 \end{bmatrix}$$

Now we just need to find one point P on the line of intersection, so that the parametric equation (in t) is $P + t(4, -19, -14)$. We can impose an additional restriction that $x = 0$ (or $y = 0$ or $z = 0$) and solve the simultaneous equations

$$2y - z = 2$$
$$-2y + 3z = 5.$$

and get $(0, 11/4, 7/2)$ is a point on the intersection. The final parametric representation is $l(t) = (0, 11/4, 7/2) + t(4, -19, -14)$.

Exercises

10. todo...

11. On the plane given by $ax + by + cz = d$, find the closest distance from $P(x_1, y_1, z_1)$ to the plane. (Hint: Refer to Example 1.44 in section 1.5)

1.6 End of Chapter Exercises

12. todo - proof based questions

Chapter 2

Linear Algebra Basics

The foundational abstraction of linear algebra is the vector space. A *vector space* is essentially a collection of objects that it makes sense to take linear combinations of. Two operations must be defined: addition and scalar multiplication. A well defined vector space meets all of the vector space axioms, which will be listed shortly. Many consequences can be drawn from these axioms, and we can build up linear algebra to solve any linear problem.

Remark: One has to specify the field of scalars (*field* just means number system) related to the vector space. In most cases here, we will be talking about real vector spaces (the field of scalars is $\mathbb{R}$). However, vector spaces can be defined with many other fields of scalars. Examples will follow the definition.

Definition 2.1 (Vector Space)

Let $\mathbb{K}$ be a field, and V be a set closed under operations addition $+ : V \times V \to V$ and multiplication $\cdot : \mathbb{K} \times V \to V$. We call V a **vector space**, or a **$\mathbb{K}$-vector space** to specify the field if the following axioms hold.

1. *(Associativity)*$x + (y + z) = (x + y) + z \ \forall x, y, z \in V$.
2. *(Commutativity)* $x + y = y + x \ \forall x, y \in V$.
3. *(Identity)* There exists some vector 0_V s.t. $x + 0_V = x \ \forall x \in V$.
4. *(Inverse)* $\forall x \in V$, $\exists y \in V$ s.t. $x + y = 0_v$.
5. *(Scalar multiplication)* $a \cdot (b \cdot x) = (a \cdot b) \cdot x \ \forall a, b \in \mathbb{K}, x \in V$.
6. *(Scalar Identity)* $1 \cdot x = x \ \forall x \in V$.
7. *(Distributivity 1)* $a \cdot (x + y) = a \cdot x + a \cdot y \ \forall a \in \mathbb{K}, x, y \in V$.
8. *(Distributivity 2)* $(a + b) \cdot x = a \cdot x + b \cdot x \ \forall a, b \in \mathbb{K}, x \in V$.

Example 2.2

The following are examples of vector spaces.

- $\mathbb{R}^n$, with addition and multiplication as we have defined so far.
- 0_V, the set containing just the zero vector, is a vector space over any field.
- The set of polynomials with degree 3 or less. Addition and multiplication are defined as $(f + g)(x) = f(x) + g(x)$ and $cf(x) = c \times f(x)$.
- $\mathbb{R}$ is a $\mathbb{Q}$-vector space, where $\mathbb{Q}$ is the set of all rationals (fractions).

Example 2.3

The following are non-examples of vector spaces.

- ϕ, the empty set is not a vector space over any field.
- The set of polynomials with degree 3 or more (using the addition and multiplication rules defined above).

More frequently, we make new vector spaces from taking subsets of existing vector spaces. For instance, there might be a subset of $\mathbb{R}^n$ that also satisfies all the axioms of being a vector space.

Definition 2.4 (Subspace)

Let V be a $\mathbb{K}$ vector space. We call $W \subseteq V$ a **subspace** of V if W forms a vector space using the inherited operations $+ : W \times W \to W$ and $\cdot : \mathbb{K} \times W \to W$.

Remark. *Because the inherited operations will automatically satisfy the axioms of a vector space, it suffices to show that (1) W is nonempty and (2) W is closed under vector addition and scalar multiplication to confirm that W is a subspace.*

The condition W is nonempty is required because the identity axiom requires a zero vector, which one obtains by multiplying 0 to an arbitrary vector in the subset.

Example 2.5

1. In $\mathbb{R}$-vector space $\mathbb{R}^3$, the set of all vectors in the form of $k\vec{i}$ forms a subspace.
2. Take $\mathbb{R}$ as a $\mathbb{Q}$-vector space. $\mathbb{Q} \subset \mathbb{R}$ is a subspace.
3. Take $\mathbb{R}$ as a $\mathbb{R}$-vector space. $\mathbb{Q} \subset \mathbb{R}$ is **not** a subspace because it is not closed under multiplication.

Proposition 2.6

Let $W, U \subseteq V$ be subspaces. The intersection $W \cap U$ is a subspace of V.

Proof. Both W and U contain 0_V, so the intersection is non-empty. We also have for $v_1, v_2 \in W \cap U$,

$$v_1 + v_2 \in U, \quad v_1 + v_2 \in W,$$

so addition is closed in $W \cap U$. Similarly, scalar multiplication closed under W and U, so is closed in the intersection. $\square$

Exercises

1. a

2.1 Span and Linear Independence

Throughout this section, let $\mathbb{K}$ be a field. We constrain ourselves to work in V, a $\mathbb{K}$-vector space.

Definition 2.7 (Span)

Let k be some positive integer. Let $\{v_1, v_2, ..., v_k\} \subseteq V$. The **span** of $\{v_1, v_2, ..., v_k\}$ is denoted as

$$\text{span}(v_1, v_2, ..., v_k)$$

and is the *smallest* subspace of V containing $\{v_1, v_2, ..., v_k\}$.

The *smallest* here means that if another subspace W contains $\{v_1, ..., v_k\}$, W cannot be a subset of $\text{span}(v_1, ..., v_k)$. How do we know such a subspace exists? We can take the intersection of all the subspaces containing $\{v_1, ..., v_k\}$

$$\text{span}(v_1, ..., v_k) = \bigcap_{W \text{ subspace containing } \{v_1,...,v_k\}} W$$

which is a subspace containing $\{v_1, ..., v_k\}$ and is a subset of all other subspaces containing $\{v_1, ..., v_k\}$. We know $V \subseteq V$ is a subspace containing $\{v_1, ..., v_k\}$, so the intersection between at least one set and is thus well-defined.

Proposition 2.8

The span of $\{v_1, v_2, ..., v_k\} \subseteq V$ is all linear combinations of the vectors in the set:

$$\text{span}\{v_1, v_2, ..., v_k\} = \{c_1v_1 + c_2v_2 + ... + c_kv_k | c_1, c_2, ...c_k \in \mathbb{K}\}$$

Proof. We first show $\text{span}\{v_1, v_2, ..., v_k\} \supseteq \{c_1v_1 + c_2v_2 + ... + c_kv_k | c_1, c_2, ...c_k \in \mathbb{K}\}$. That is, for every W subset containing $v_1, ..., v_k$, W must also contain $c_1v_1 + ... + c_kv_k$.
We now show $\text{span}\{v_1, v_2, ..., v_k\} \subseteq \{c_1v_1 + c_2v_2 + ... + c_kv_k | c_1, c_2, ...c_k \in \mathbb{K}\}$. The right side is a set that nonempty, is closed under addition and scalar multiplication, and contains $v_1 = 1v_1 + 0v_2 + ... + 0v_k$,...,$v_k = 0v_1 + ... + 0v_{k-1} + 1v_k$. Therefore, it is one of the W subspaces whose intersection is used to construct the span. □

It can be difficult to how to think about what the span of a set of vectors looks like, though it is also important to develop an intuition for it as more complex techniques are developed. It is also important to consider what vectors span a given subspace.

Example 2.9

For example, the vector space of $\mathbb{R}^3$ is spanned by the set of unit vectors $\{i, j, k\}$.

Vector spaces' sizes can be compared by the minimum number of vectors required to span them. This "dimensionality" will be used more later, but for now it is important that vector spaces can be either finite or infinite dimensional.

Example 2.10

Vector spaces like $\mathbb{R}^n$ can be spanned by n vectors and thus are finite dimensional. Vector spaces that are composed of less linear objects, like the set of all functions $\mathbb{R} \to \mathbb{R}$, are often infinite dimensional. In this case, note that there's no finite list of functions such that all other functions from $\mathbb{R} \to \mathbb{R}$ are linear combinations of them.