Vectors in $\mathbb{R}^n$

Math 1410 Linear Algebra

A note on software

Visualizing objects (vectors, lines, planes, etc.) in 3 dimensions can be tricky on paper. Fortunately, these days we're equipped with all kinds of computer software that helps us with these visualizations. Some software, like Mathematica or Maple, is very good but also very expensive. (You might check to see what's available in the computer labs on campus.) There are also lots of free options, including many interactive websites. You might find that simple Google searches (like "3D vector visualization") turn up lots of options.

I'll be using software in class called Geogebra. If you want to try it yourself, it's free, open-source software available for Windows, OSX, and Linux. You can download it at geogebra.org.

Cartesian coordinates

- Our “base” number system is the real numbers $\mathbb{R}$.
- Definition of $\mathbb{R}$ complicated - includes rational (fraction) and irrational numbers.
- Visualize as “number line” – geometrically one-dimensional.
- Define $\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$ – set of ordered pairs of real numbers.
- Cartesian plane visualizes $\mathbb{R}^2$ using two “coordinate axes”.

The set $\mathbb{R}^n$

We can extend Descartes' construction to define the set

$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) \,|\, x_i \in \mathbb{R}, i = 1, \ldots, n\}.$$

This is often referred to as "n-dimensional Euclidean space". (The adjective Euclidean refers to the geometric structure of lines and planes.)
We can only visualize $\mathbb{R}^n$ for $n = 1, 2$, or 3.

Distance in $\mathbb{R}^n$

The distance between two points $x, y \in \mathbb{R}$ is given by

$$d(x, y) = |x - y|,$$

where $|a|$ denotes the absolute value of $a \in \mathbb{R}$. Notice that $|x - y| = \sqrt{(x - y)^2}$.

In $\mathbb{R}^2$, distance is given by the Pythagorean Theorem:

$$d((x_1, y_1), (x_2, y_2)) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

We can prove that the same pattern holds in $\mathbb{R}^3$:

$$d((x_1, y_1, z_1), (x_2, y_2, z_2)) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}.$$

We define distance in $\mathbb{R}^n$ for $n \geq 4$ by insisting that this pattern continues.

Vectors, geometric and algebraic

Our interest in Chapter 4 is the study of vectors. The meaning of the term "vector" varies depending on the context:

- Geometry/physics: a quantity with both magnitude and direction. (Basically, an arrow.) Think of velocity, force, etc.
- Algebra: quantities that can be added together, and multiplied by numbers (scalars), subject to certain rules.
- Data analytics: ordered arrays of information that can be manipulated.
- And so on...

The third point above isn't all that different from the second. We'll be interested in seeing how the second point is connected to the first.

Geometric vectors

A geometric vector in $\mathbb{R}^n$ is visualized as an arrow. We if the "tail" of our arrow/vector $\vec{v}$ is at a point $P = (x_1, \ldots x_n)$ and the "tip" is at a point $Q = (y_1, \ldots, y_n)$, we write $\vec{v} = \overrightarrow{PQ}$.
Viewed geometrically, a vector is determined by its length (magnitude) and direction, but not its *location*. Thus, if $R = (x_1 + a_1, \ldots, x_n + a_n)$ and $S = (y_1 + a_1, \ldots, y_n + a_n)$ for some constants $a_1, \ldots, a_n$, then $\overrightarrow{RS} = \overrightarrow{PQ}$.

Example

Numerical representation of vectors

Let $\vec{v} = \overrightarrow{PQ}$ as on the previous slide. Note that all the information about $\vec{v}$ is contained in the differences $y_1 - x_1, y_2 - x_2, \ldots, y_n - x_n$. We represent $\vec{v}$ numerically by

$$\vec{v} = \langle y_1 - x_1, y_2 - x_2, \ldots, y_n - x_n \rangle.$$

Notice that this representation does not depend on the location of $\vec{v}$:

Position vectors

Since the location of a geometric vector doesn't matter, it's often convenient to locate the tail of the vector at the origin $O = (0, 0, \ldots, 0)$. If $P = (x_1, x_2, \ldots, x_n)$ is any other point in $\mathbb{R}^n$, we have the corresponding position vector

$$\vec{p} = \overrightarrow{OP} = \langle x_1, x_2, \ldots, x_n \rangle.$$

Note: there is a "one-to-one" correspondence between points in $\mathbb{R}^n$ and vectors in $\mathbb{R}^n$. For every point P we have the corresponding position vector $\vec{p}$, and vice versa.

Length of a vector

The length, or magnitude of a vector $\vec{v}$ is denoted by $\|\vec{v}\|$. If $\vec{v} = \overrightarrow{PQ}$, then $\|\vec{v}\|$ is simply the distance from P (the tail of $\vec{v}$) to Q (the tip). Thus,

$$\|\vec{v}\| = \sqrt{v_1^2 + \cdots + v_n^2} = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2},$$

where the numbers $v_i = x_i - y_i$, $i = 1, 2, \ldots, n$, are called the components of $\vec{v}$.

Addition of geometric vectors

Geometric vectors are added according to the "tip-to-tail" or "parallelogram" rule. We can sketch the procedure as follows:

Scalar multiplication of geometric vectors

Let $c \in \mathbb{R}$ be a real number, and let $\vec{v}$ be a vector in $\mathbb{R}^n$. The scalar multiple $c\vec{v}$ is defined as follows:

- If $c > 0$, then $c\vec{v}$ points in the same direction as $\vec{v}$, and has length $\|c\vec{v}\| = c\|\vec{v}\|$.
- If $c = 0$, then $c\vec{v}$ is the "zero vector" $\vec{0}$.
- If $c < 0$, then $c\vec{v}$ points in the opposite direction as $\vec{v}$, and has length $\|c\vec{v}\| = (-c)\|\vec{v}\|$.

Notes:

1. We can sum up all of the above with the rule $\|c\vec{v}\| = |c|\|\vec{v}\|$.
2. The zero vector is a little weird. It has magnitude zero, but no direction.

Algebraic vectors

We've already met algebraic vectors. Let's let $\mathbb{R}^{n,1}$ denote the set of all $n \times 1$ column vectors

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

We define

$$\vec{x} + \vec{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix} \text{ and } c\vec{x} = c \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix},$$

where c is a scalar. (This was our definition for matrices in general.)

Examples

Properties of vector algebra

The addition and scalar multiplication of vectors follows the following rules:
For any vectors $\vec{x}, \vec{y}, \vec{z}$ and scalars $c, d \in \mathbb{R}$,

1. $\vec{x} + \vec{y} = \vec{y} + \vec{x}$
2. $\vec{x} + (\vec{y} + \vec{z}) = (\vec{x} + \vec{y}) + \vec{z}$
3. $\vec{x} + \vec{0} = \vec{x}$, where $\vec{0} = \begin{bmatrix} 0 & 0 & \cdots & 0 \end{bmatrix}^T$
4. $\vec{x} + (-\vec{x}) = \vec{0}$, where $-\vec{x} = \begin{bmatrix} -x_1 & -x_2 & \cdots & -x_n \end{bmatrix}^T$
5. $1 \cdot \vec{x} = \vec{x}$
6. $c(d\vec{x}) = (cd)\vec{x}$
7. $c(\vec{x} + \vec{y}) = c\vec{x} + c\vec{y}$
8. $(c + d)\vec{x} = c\vec{x} + d\vec{x}$.

Vector spaces

Whenever we have a set of objects on which we can define an addition and scalar multiplication satisfying the same 8 rules as vectors in $\mathbb{R}^n$, we call that set a vector space.

There are lots of examples: $\mathbb{R}^n$, of course, but also the space of all $m \times n$ matrices, and the space of all polynomials of degree less than or equal to k for some k. (These are "finite-dimensional" vectors spaces.)

Other examples include the space of all real-valued functions defined on some interval $[a, b]$, and the space of all sequences $(a_1, a_2, a_3, \ldots)$. (These are *infinite-dimensional* vector spaces.)

For Math 1410, the only vector spaces are $\mathbb{R}^n$, for $n = 1, 2, 3 \ldots$.
For general vector spaces, take Math 3410.

Algebraic vs. Geometric

We've defined vectors in $\mathbb{R}^n$ in two ways:

- Geometric vectors $\vec{v} = \overrightarrow{PQ} = \langle y_1 - x_1, \ldots, y_n - x_n \rangle$, where $P = (x_1, \ldots, x_n)$ is the point at the tip of the vector, and $Q = (y_1, \ldots, y_n)$ is the point at the tail.
- Algebraic vectors $\vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$, where $v_1, \ldots, v_n \in \mathbb{R}$ are real numbers.

Goal: adding geometric vectors is the same as adding algebraic vectors, and the same is true for scalar multiplication.

Examples

Linear combinations

One of the most fundamental concepts in Linear Algebra is the linear combination.
Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ be vectors in $\mathbb{R}^n$ and let $c_1, c_2, \ldots, c_k$ be scalars.
A vector $\vec{v}$ is a linear combination of the vectors $\vec{v}_i$ if

$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k$$

Algebraic examples

Geometric examples

Another example

Problem: given $\vec{u} = \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}$, find the following points:

1. P, half-way between the tip of $\vec{u}$ and the tip of $u \vec{+} v$.
2. Q, half-way between the tip of $\vec{v}$ and the tip of $u \vec{+} v$.

Summary of results so far

Recall: Last week we studied the algebraic and geometric properties of vectors in $\mathbb{R}^n$. We had three basic objects:

1. Points in $\mathbb{R}^n$: $P = (x_1, x_2, \ldots, x_n)$
2. Geometric vectors in $\mathbb{R}^n$: $\vec{v} = \langle x_1, x_2, \ldots, x_n \rangle$
3. Algebraic vectors in $\mathbb{R}^n$: $\vec{v} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

Note that all three concepts provide the same information:

- We can identify points P with position vectors $\vec{p}$
- Component-wise (algebraic) addition of vectors corresponds to "tip-to-tail" (geometric) addition.
- Component-wise (algebraic) scalar multiplication also agrees with geometric scalar multiplication.

Linear independence

We say that a set of vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ is linearly dependent if one of the vectors can be written as a linear combination of the other vectors; for example, if

$$\vec{v}_1 = c_2\vec{v}_2 + \cdots + c_k\vec{v}_k$$

for scalars $c_2, \ldots, c_k$.
This tells us we don't really need the first vector, since we can get it from the other ones.
On the other hand, our vectors are linearly independent if the above is impossible. Equivalently, the only way we can have

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \vec{0}$$

is if $c_1 = 0, c_2 = 0, \ldots, c_k = 0$.

Example

Example

Span

The span of vectors $\vec{v}_1, \ldots, \vec{v}_k$ in $\mathbb{R}^n$ is the set of all possible linear combinations we can form from these vectors. Thus,

$$\vec{w} \in \text{span}\{\vec{v}_1, \ldots, \vec{v}_k\}$$

if and only if

$$\vec{w} = c_1 \vec{v}_1 + \cdots + c_k \vec{v}_k$$

for scalars $c_1, \ldots, c_k$.

Example

Let $\vec{u} = \begin{bmatrix} 2 & -1 & 4 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} -1 & 3 & 1 \end{bmatrix}$. Describe:

1. $\text{span}\{\vec{u}\}$
2. $\text{span}\{\vec{u}, \vec{v}\}$

Subspaces

A subspace of $\mathbb{R}^n$ is a subset $U \subseteq \mathbb{R}^n$ that satisfies all of the 8 vector space properties described earlier. One can show that U is a subspace if:

1. $\vec{0}$ belongs to U.
2. If $\vec{u}_1, \ldots, \vec{u}_k$ are vectors in U, then any linear combination $c_1\vec{u}_1 + \cdots + c_k\vec{u}_k$ also belongs to U.

Consequence: every subspace can be written as the span of some vectors.

Lines in $\mathbb{R}^2$

We know many ways to describe a line in the plane: two points, point and slope, $y = mx + b$, $y - y_0 = m(x - x_0)$, etc.

Lines in $\mathbb{R}^3$ - I

In three dimensions, two points still determine a line, but "slope" has no meaning.
However, given points P and Q, we have a direction: the vector $\overrightarrow{PQ}$. This leads to a vector equation of a line:

Example

Find the line in $\mathbb{R}^3$ through the points $P = (0, 1, 3)$ and $Q = (2, -1, 4)$.

Lines in $\mathbb{R}^3$ - II

Note we have various ways of describing a line again: two points, point and direction, or a vector equation. We often are also interested in parametric equations of a line.

Lines as subspaces

Remark: If $\vec{v}$ is a nonzero vector in $\mathbb{R}^n$, then

$$U = \text{span}\{\vec{v}\} = \{t\vec{v} \mid t \in \mathbb{R}\}$$

is a line through the origin. It is also a subspace. In fact, all "one-dimensional" subspaces of $\mathbb{R}^n$ are of this form.

Dot products

The dot product (or scalar product) of two vectors $\vec{u} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}^T$, $\vec{v} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}^T$ is defined algebraically by

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{i=1}^{n} u_i v_i.$$

If we view $\vec{u}$ and $\vec{v}$ as geometric vectors, then we define

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

where θ is the angle between the two vectors.

Examples

Algebraic definition = geometric definition

Theorem

For any vectors $\vec{u}, \vec{v}$ in $\mathbb{R}^n$, the geometric and algebraic definitions of the dot product coincide. That is,

$$\vec{u} \cdot \vec{v} = u_1 v_1 + \cdots + u_n v_n = \|\vec{u}\| \|\vec{v}\| \cos\theta.$$

The proof uses the law of cosines: for any triangle with sides of length a, b, c, if θ is the interior angle opposite the side of length c, then

$$c^2 = a^2 + b^2 - ab\cos\theta.$$

Properties of the dot product

Theorem

For any vectors $\vec{u}, \vec{v}, \vec{w}$ and any scalars $a, b \in \mathbb{R}$, we have:

1. $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{w}$
2. $\vec{u} \cdot (a\vec{v} + b\vec{w}) = a\vec{u} \cdot \vec{v} + b\vec{u} \cdot \vec{w}$.
3. *If $\vec{u} \cdot \vec{v} = 0$ for all vectors $\vec{v}$, then $\vec{u} = 0$.*
4. $\vec{u} \cdot \vec{u} = \|\vec{u}\|^2$
5. $\vec{u} \cdot \vec{u} \geq 0$, *and $\vec{u} \cdot \vec{u} = 0$ if and only if $\vec{u} = \vec{0}$*
6. $|\vec{u} \cdot \vec{v}| \leq \|\vec{u}\|\|\vec{v}\|$ *(Cauchy-Schwarz inequality)*

Some proofs

Examples

Triangle inequality

A consequence of the Cauchy-Schwarz inequality is the triangle inequality:

Theorem

For any vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$, we have

$$\|\vec{u}+\vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|.$$

Another result that is often useful is the "reverse triangle inequality":

$$|\|\vec{u}\| - \|\vec{v}\|| \leq \|\vec{u}-\vec{v}\|.$$

Dot product and geometry

We saw above that $\vec{u} \cdot \vec{v} = \|\vec{u}\| \|\vec{v}\| \cos\theta$ for any vectors $\vec{u}$ and $\vec{v}$. If we know the components of two non-zero vectors $\vec{u}$ and $\vec{v}$, then we can find the angle between them:

$$\cos\theta = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \|\vec{v}\|}.$$

In particular, note that $\vec{u} \cdot \vec{v} = 0$ if and only if $\cos\theta = 0$; i.e., $\theta = \dfrac{\pi}{2}$.

Definition

We say that two vectors $\vec{u}, \vec{v}$ in $\mathbb{R}^n$ are orthogonal if $\vec{u} \cdot \vec{v} = 0$.

Example

Find the angle between the vectors $\vec{u} = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} -1 \\ 4 \\ -2 \end{bmatrix}$.

Example

Two lines L_1 and L_2 intersect at the point $P = (1, 0, -2)$. If L_1 also passes through the point $Q = (3, 5, 1)$ and L_2 passes through the point $R = (-2, 1, 3)$, find the angle between the two lines.

Unit vectors

A unit vector is a vector of length 1. Note that if $\vec{v} \neq \vec{0}$, one possible unit vector is

$$\hat{v} = \frac{1}{\|\vec{v}\|}\vec{v}.$$

The standard unit vectors in $\mathbb{R}^n$ are the vectors $\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \cdots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$. In $\mathbb{R}^3$, common notation is

$$\hat{\imath} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \hat{\jmath} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \hat{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Components

The dot product lets us tell "how much" of a vector $\vec{v}$ lies in the same direction as a vector $\vec{u}$ (i.e. how closely the two vectors are aligned).

Definition

Let $\vec{u}$ be a nonzero vector. The component of $\vec{v}$ in the direction of $\vec{u}$ is defined by

$$\text{comp}_{\vec{u}}\,\vec{v} = \frac{\vec{v}\cdot\vec{u}}{\|\vec{u}\|} = \|\vec{v}\|\cos\theta.$$

Note: we divide by $\|\vec{u}\|$ since we only care about the direction of $\vec{u}$, not its length.
Remark: Note that components of a vector $\vec{v}$ in the direction of the standard unit vectors are just the usual components of $\vec{v}$.

Projections

Given vectors $\vec{u}, \vec{v}$, note that $\text{comp}_{\vec{u}}\, \vec{v}$ is a number. The corresponding vector quantity is called the projection of $\vec{v}$ onto $\vec{u}$:

Definition

Let $\vec{u}$ be a nonzero vector. The projection of $\vec{v}$ onto $\vec{u}$ is the vector $\text{proj}_{\vec{u}}\, \vec{v}$ defined by

$$\text{proj}_{\vec{u}}\, \vec{v} = \left(\frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|^2}\right) \vec{u} = (\text{comp}_{\vec{u}}\, \vec{v})\hat{u},$$

where $\hat{u} = \frac{1}{\|\vec{u}\|}\vec{u}$ is the unit vector in the direction of $\vec{u}$.

Note that $\text{proj}_{\vec{u}}\, \vec{v}$ is the vector in the direction of $\vec{u}$ whose length is $\text{comp}_{\vec{u}}\, \vec{v}$.

Orthogonal decomposition

Theorem

Let $\vec{u}$ be a nonzero vector in $\mathbb{R}^n$. Then given any vector $\vec{v}$ in $\mathbb{R}^n$, there exist vectors $\vec{v}_{||}$ and $\vec{v}_{\perp}$ such that:

1. *$\vec{v}_{||}$ is parallel to $\vec{u}$,*
2. *$\vec{v}_{\perp}$ is orthogonal to $\vec{u}$, and*
3. *$\vec{v} = \vec{v}_{||} + \vec{v}_{\perp}$.*

Examples

Let $\vec{u} = \begin{bmatrix} 1 & -2 & 2 \end{bmatrix}^T$ and $\vec{v} = \begin{bmatrix} 3 & 4 & -1 \end{bmatrix}^T$. Find:

1. $\operatorname{comp}_{\vec{u}} \vec{v}$
2. $\operatorname{proj}_{\vec{u}} \vec{v}$
3. $\vec{v}_{||}$ and $\vec{v}_{\perp}$.

Example

Let L be the line given by $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix} + t \begin{bmatrix} 4 \\ 0 \\ -3 \end{bmatrix}$. Find the point on L that is closest to the point $P = (5, -3, 7)$.

Planes in $\mathbb{R}^3$

A plane in $\mathbb{R}^3$ is a flat, two-dimensional surface sitting inside of $\mathbb{R}^3$. (Think of a copy of $\mathbb{R}^2$ inside $\mathbb{R}^3$.) The most basic examples are the coordinate planes:

- The xy-plane: $\{(x, y, 0) \mid x, y \in \mathbb{R}\}$, or $z = 0$.
- The yz-plane: $\{(0, y, z) \mid y, z \in \mathbb{R}\}$, or $x = 0$.
- The xz-plane: $\{(x, 0, z) \mid x, z \in \mathbb{R}\}$, or $y = 0$.

More general planes can be defined in many ways:

- As a single linear equation: $ax + by + cz = d$.
- By specifying three points (not all on the same line).
- By giving two intersecting lines.
- As the span of two vectors.
- By giving a point and a normal vector

Scalar equation

Choose a point $P_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$ and a non-zero vector $\vec{n} = \begin{bmatrix} a & b & c \end{bmatrix}^T$.

The plane determined by P_0 and $\vec{n}$ is the set of all points P such that the vector $\overrightarrow{P_0P}$ is perpendicular to $\vec{n}$.

Example

Determine the plane through the point $P_0 = (2, -3, 1)$ with normal vector $\vec{n} = \langle 5, -2, 3 \rangle$.

Example

Find the shortest distance from the point $Q = (2, 1, -3)$ to the plane with equation $3x - y + 4z = 1$.

Example

Describe the intersection of the planes given by $2x - y + 4z = 2$ and $x + 2y - z = 5$.

Planes as spans

Recall that given vectors $\vec{u}$ and $\vec{v}$, we define

$$\operatorname{span}\{\vec{u}, \vec{v}\} = \{a\vec{u} + b\vec{v} \mid a, b \in \mathbb{R}\}.$$

Suppose we're given a plane $a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$ in $\mathbb{R}^3$. Think of this single equation as a system of one equation in three variables.
How would we have handled this in Chapter 1?

Planes through $(0, 0, 0)$ are subspaces

A plane through the origin is given by $ax + by + cz = 0$. This is equivalent to the vector equation

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = s \begin{bmatrix} -b \\ a \\ 0 \end{bmatrix} + t \begin{bmatrix} -c \\ 0 \\ a \end{bmatrix}$$

Returning to the scalar equation

Suppose we are given a plane through $P_0 = (x_0, y_0, z_0)$ such that the vectors $\vec{u} = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}^T$ and $\vec{v} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}^T$ are parallel to the plane. How can we recover the scalar equation?

The cross product

There are two ways to define the cross product $\vec{u} \times \vec{v}$ of two vectors in $\mathbb{R}^3$:

- Geometrically: If $\vec{u}$ and $\vec{v}$ are parallel, we set $\vec{u} \times \vec{v} = \vec{0}$. Given non-parallel vectors $\vec{u}$ and $\vec{v}$, the cross product $\vec{u} \times \vec{v}$ satisfies:
  - $\vec{u} \times \vec{v}$ is orthogonal to both $\vec{u}$ and $\vec{v}$
  - $\|\vec{u} \times \vec{v}\| = \|\vec{u}\|\|\vec{v}\| \sin\theta$, where θ is the angle between $\vec{u}$ and $\vec{v}$.
  - The direction of $\vec{u} \times \vec{v}$ satisfies the "right-hand rule".
- Algebraically: Given $\vec{u} = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}^T$ and $\vec{v} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}^T$, we define

$$\vec{u} \times \vec{v} = \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - v_3 u_1 \\ u_1 v_2 - u_2 v_1 \end{bmatrix}.$$

Example

Deriving the cross product

Recall: given $\vec{u}$ and $\vec{v}$, we want $\vec{n} = \vec{u} \times \vec{v}$ to satisfy $\vec{n} \cdot \vec{u} = 0$ and $\vec{n} \cdot \vec{v} = 0$. Assume

$$\vec{u} = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}^T, \vec{u} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}^T, \vec{n} = \begin{bmatrix} a & b & c \end{bmatrix}^T.$$

Then we get a system of two equations in three variables:

$$\begin{array}{ccccccc} u_1 a & + & u_2 b & + & u_3 c & = & 0 \\ v_1 a & + & v_2 b & + & v_3 c & = & 0 \end{array}$$

Remembering the cross product I

Note that the x-component of $\vec{u} \times \vec{v}$ is the 2×2 determinants of the y and z components of $\vec{u}$ and $\vec{v}$, and so on. We can remember all of this as a single 3×3 "determinant":

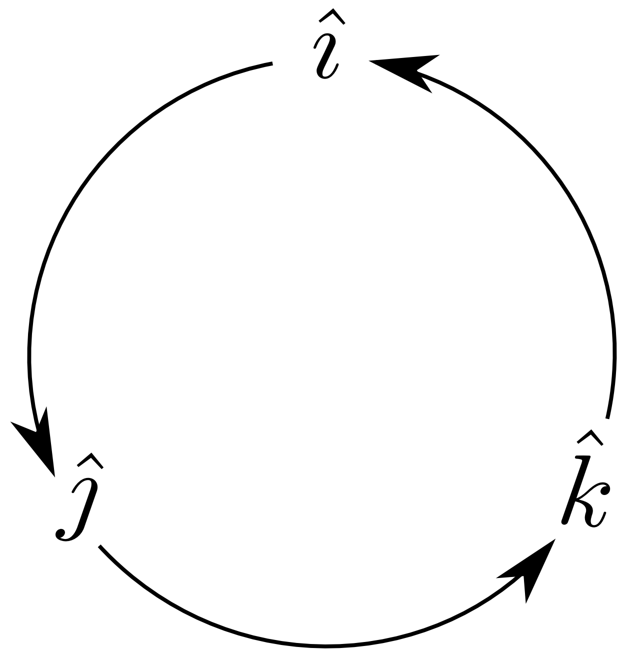
$$\vec{u} \times \vec{v} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} \hat{\imath} - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} \hat{\jmath} + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \hat{k}$$

Remembering the cross product II

Another way to compute the cross product is to work in terms of the basic unit vectors $\hat{\imath}, \hat{\jmath}$, and $\hat{k}$. It's easy to check (exercise) that

$$\hat{\imath} \times \hat{\jmath} = \hat{k}, \quad \hat{\jmath} \times \hat{k} = \hat{\imath}, \quad \hat{k} \times \hat{\imath} = \hat{\jmath},$$

which we can remember like so:



Properties of the cross product

Theorem

If $\vec{u} = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}^T, \vec{v} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}^T$, *and* $\vec{w} = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix}^T$, *then* $\vec{u} \cdot (\vec{v} \times \vec{w}) = \det \begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix}$.

The expression $\vec{u} \cdot (\vec{v} \times \vec{w})$ is often called the scalar triple product.
From the above theorem, we can deduce the following:

- $\vec{u} \times \vec{v}$ is a vector orthogonal to both $\vec{u}$ and $\vec{v}$.
- $\vec{u} \times \vec{0} = \vec{0} \times \vec{u} = \vec{0}$.
- If $\vec{u}$ and $\vec{v}$ are parallel, then $\vec{u} \times \vec{v} = \vec{0}$.
- $\vec{u} \times \vec{v} = -(\vec{v} \times \vec{u})$.
- $(c\vec{u}) \times \vec{v} = \vec{u} \times (c\vec{v}) = c(\vec{u} \times \vec{v})$ for any $c \in \mathbb{R}$.
- $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$.

The Lagrange Identity

The following result, called the Lagrange identity, relates the cross product to the dot product:

Theorem

For any two vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^3$, we have

$$\|\vec{u} \times \vec{v}\|^2 = \|\vec{u}\|^2\|\vec{v}\|^2 - (\vec{u} \cdot \vec{v})^2.$$

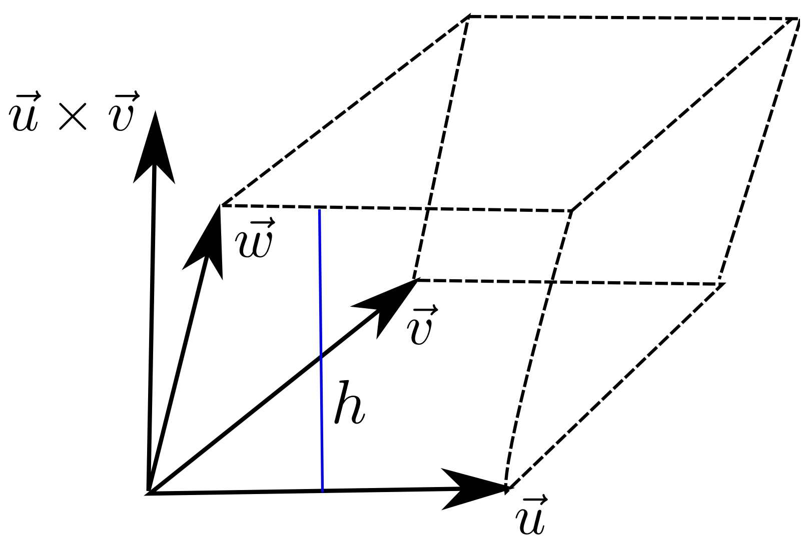
Consequence: $\|\vec{u} \times \vec{v}\| = \|\vec{u}\|\|\vec{v}\| \sin\theta$, which is the area of the parallelogram spanned by $\vec{u}$ and $\vec{v}$.

Example

Find the area of the triangle with vertices $P = (2, 1, 0)$, $Q = (3, -1, 1)$, and $R = (1, 0, 1)$.

Volumes

We mentioned in our discussion of determinants that a three-by-three determinant is related to volume. We can see this using the scalar triple product: the vectors $\vec{u}, \vec{v}, \vec{w}$ span a parallelepiped, like so:



The volume is given by $V = Ah$, where A is the area of the base, and h is the height.

Example

Show that the shortest distance from a point P to the line L through the point P_0 with direction vector $\vec{d}$ is

$$\frac{||\overrightarrow{P_0P} \times \vec{d}||}{||\vec{d}||}.$$

Example

Find the equation of the plane that contains the lines

$$L_1 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$$

$$L_2 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} + t \begin{bmatrix} 3 \\ -4 \\ 2 \end{bmatrix}$$

Example

Find the equation of the plane containing the points $P = (0, 1, 2)$, $Q = (-2, 3, 5)$, and $R = (1, 1, -3)$.

Example

Find the distance between the following skew lines, and find the points on each line that are closest together:

$$\begin{bmatrix} x & y & z \end{bmatrix}^T = \begin{bmatrix} 3 & 0 & 1 \end{bmatrix}^T + s\begin{bmatrix} 2 & 1 & -3 \end{bmatrix}^T$$
$$\begin{bmatrix} x & y & z \end{bmatrix}^T = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix}^T + t\begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^T$$

Subspaces

Recall: a subspace $V \subseteq \mathbb{R}^n$ is a set of vectors in $\mathbb{R}^n$ that contains $\vec{0}$, such that if $\vec{x}, \vec{y} \in V$, then $a\vec{x} + b\vec{y} \in V$ for any scalars a and b.

Example

The set $V = \{(x, y, 2x - 3y) \mid x, y \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^3$.

Example

Let $\vec{x} = \begin{bmatrix} 2 & -3 & 4 \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} -1 & 0 & 3 \end{bmatrix}$ be two vectors in $\mathbb{R}^3$. Then

$$\operatorname{span}\{\vec{x}, \vec{y}\} = \{a\vec{x} + b\vec{y} \mid a, b \in \mathbb{R}\}$$

is a subspace of $\mathbb{R}^3$.

Spans are subspaces

In general, let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ be vectors in $\mathbb{R}^n$. Then

$$V = \operatorname{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$$

is a subspace of $\mathbb{R}^n$.

Basis for a subspace

One of the most important concepts in linear algebra is that of a basis.

Definition

Let V be a subspace of $\mathbb{R}^n$. We say that a set of vectors

$$B = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$$

is a basis for V if span $B = V$, **and** the vectors in B are linearly independent.

Example

Find a basis for the subspace of $\mathbb{R}^3$ defined by the equation $2x - y + 3x = 0$.

Example

Find a basis for the subspace

$$V = \{(2x - y, x + 2y, x, -x + 4z) \,|\, x, y \in \mathbb{R}\} \subseteq \mathbb{R}^4$$

Orthogonal sets of vectors

Recall that two vectors $\vec{u}$ and $\vec{v}$ are orthogonal if $\vec{u} \cdot \vec{v} = 0$.

Definition

We say that a set $S = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is an orthogonal set of vectors if:

- $\vec{v}_i \neq \vec{0}$, for all $i = 1, 2, \ldots, k$,
- $\vec{v}_i \cdot \vec{v}_j = 0$ for all $i \neq j$.

Example

Show that the set $\left\{ \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} \right\}$ is an orthogonal set of vectors.

Orthogonal sets are linearly independent

One reason that orthogonal sets of vectors are useful is that they're automatically linearly independent:

Theorem

Let $S = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ *be an orthogonal set of vectors. Then the vectors in* S *are linearly independent.*

Consequence: If the vectors in S span a subspace V, S is automatically a basis.

Fourier and Pythagorus

There are two important results involving orthogonal sets of vectors:

Theorem (Pythogorean Theorem)

For any set $S = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ *of orthogonal vectors, we have*

$$\|\vec{v}_1 + \vec{v}_2 + \cdots + \vec{v}_k\|^2 = \|\vec{v}_1\|^2 + \|\vec{v}_2\|^2 + \cdots + \|\vec{v}_k\|^2.$$

Theorem (Fourier Expansion Theorem)

Let $S = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ *be an orthogonal set of vectors. For any* $\vec{v} \in \operatorname{span} S$*, we have*

$$\vec{v} = \left(\frac{\vec{v}_1 \cdot \vec{v}}{\|\vec{v}_1\|^2}\right)\vec{v}_1 + \left(\frac{\vec{v}_2 \cdot \vec{v}}{\|\vec{v}_2\|^2}\right)\vec{v}_2 + \cdots + \left(\frac{\vec{v}_k \cdot \vec{v}}{\|\vec{v}_k\|^2}\right)\vec{v}_k$$

Orthonormal sets of vectors

Recall that a vector $\vec{v}$ is a unit vector if $\|\vec{v}\| = 1$. For any vector $\vec{u} \neq \vec{0}$, we know that

$$\hat{u} = \frac{1}{\|\vec{u}\|}\vec{u}$$

is a unit vector. (We often say that $\hat{u}$ is normalized.)

Definition

We say that a set of vectors $S = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is an orthonormal set of vectors if

$$\vec{v}_i \cdot \vec{v}_j = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}.$$

Orthonormal bases

Definition

Let $V \subseteq \mathbb{R}^n$ be a subspace. We say that $B = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is an orthonormal basis (ONB) for V if $\operatorname{span} B = V$ and B is an orthonormal set of vectors.

Note: Suppose B as above is an orthonormal basis for a subspace V. If $v \in V$, then

$$v = x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_k\vec{v}_k$$

for scalars $x_1, x_2, \ldots, x_k$. In general, finding the values of these scalars requires solving a system of n equations in k variables. For an ONB, it's easy:

Orthogonal matrices

Definition

An $n \times n$ matrix A is said to be orthogonal if $A^T = A^{-1}$. That is, A satisfies $AA^T = A^T A = I_n$.

Theorem

An $n \times n$ matrix A is orthogonal if and only if the columns of A form an orthonormal basis of $\mathbb{R}^n$

Proof.

Let $A = [\vec{v}_1|\vec{v}_2|\cdots|\vec{v}_n]$. The (i,j)-entry of $A^T A$ is then $\vec{v}_i \cdot \vec{v}_j$. $\square$

Example

Show that the matrix $A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$ is orthogonal.

Example

Let $V \subseteq \mathbb{R}^3$ be the subspace spanned by the vectors $\vec{v}_1 = \begin{bmatrix} 3 & 0 & -4 \end{bmatrix}^T$ and $\vec{v}_2 = \begin{bmatrix} 1 & -1 & 2 \end{bmatrix}^T$. Determine if the vector $\vec{v} = \begin{bmatrix} 3 & -4 & 2 \end{bmatrix}^T$ belongs to V.

Solution 1: Solve the system:

Solution 2: Find the normal vector (note V is a plane).

Solution 3: Find an orthogonal basis.

Gram-Schmidt

The Gram-Schmidt Procedure is an algorithm that converts a given basis into an orthonormal basis. The procedure is as follows: Let $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ be any basis for a subspace $V \subseteq \mathbb{R}^n$. We define new vectors $\vec{u}_1, \ldots, \vec{u}_k$ as follows:

$$\begin{aligned}
\vec{u}_1 &= \vec{v}_1. \\
\vec{u}_2 &= \vec{v}_2 - \operatorname{proj}_{\vec{u}_1} \vec{v}_2 \\
\vec{u}_3 &= \vec{v}_3 - \operatorname{proj}_{\vec{u}_1} \vec{v}_3 - \operatorname{proj}_{\vec{u}_2} \vec{v}_3 \\
&\vdots \quad \vdots \\
\vec{u}_k &= \vec{v}_k - \operatorname{proj}_{\vec{u}_1} \vec{v}_k - \operatorname{proj}_{\vec{u}_2} \vec{v}_k - \cdots - \operatorname{proj}_{\vec{u}_{k-1}} \vec{v}_k
\end{aligned}$$

Then we set $\hat{u}_j = \dfrac{1}{\|\vec{u}_j\|}\vec{u}_j$ for $j = 1, \ldots, n$, and the result is an orthonormal set with the same span as B.

Example

Suppose a subspace $V \subseteq \mathbb{R}^4$ has the basis $B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \right\}$. Use the Gram-Schmidt procedure to obtain an orthonormal basis.

Orthogonal projection

Recall: let $V \subseteq \mathbb{R}^n$ be a subspace with orthogonal basis $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$. If $v \in V$, then we have the Fourier expansion

$$\vec{v} = \left(\frac{\vec{v}_1 \cdot \vec{v}}{\|\vec{v}_1\|^2}\right)\vec{v}_1 + \left(\frac{\vec{v}_2 \cdot \vec{v}}{\|\vec{v}_2\|^2}\right)\vec{v}_2 + \cdots + \left(\frac{\vec{v}_k \cdot \vec{v}}{\|\vec{v}_k\|^2}\right)\vec{v}_k.$$

What if $\vec{v} \notin V$? If that's the case, then the right-hand side above can't equal $\vec{v}$ (since it belongs to V). What we get instead is the orthogonal projection $\operatorname{proj}_V \vec{v}$ of $\vec{v}$ onto V.

Theorem

Let $\vec{w} = \operatorname{proj}_V \vec{u}$ be the orthogonal projection of $\vec{u}$ onto a subspace V with respect to an orthogonal basis. Then $\vec{w}$ is the closest point in V to $\vec{u}$. (That is, $\|\vec{w} - \vec{u}\|$ is as small as possible.)

Example

Let $V \subseteq \mathbb{R}^3$ be the plane spanned by the vectors $\vec{v} = \begin{bmatrix} 0 & -3 & 4 \end{bmatrix}^T$ and $\vec{w} = \begin{bmatrix} 2 & -1 & 2 \end{bmatrix}$. Find the point in V that is closest to the point $P = (2, 0, 4)$.

Orthogonal complement

Let $V \subseteq \mathbb{R}^n$ be a subspace. We define the orthogonal complement $V^\perp$ of V to be the set

$$V^\perp = \{\vec{w} \in \mathbb{R}^n \,|\, \vec{w} \cdot \vec{v} = 0 \text{ for all } \vec{v} \in V\}.$$

Theorem

If V is a subspace, then $V^\perp$ is also a subspace. Moreover, every $\vec{x} \in \mathbb{R}^n$ can be written uniquely in the form $\vec{x} = \vec{v} + \vec{w}$, where $\vec{v} \in V$ and $\vec{w} \in V^\perp$.

Finding $V^\perp$

Given a subspace V, how do we find $V^\perp$? First, suppose $B = \{\vec{v}_1, \ldots, \vec{v}_k\}$ is a basis for V. If $\vec{w} \in V^\perp$, then we must have

$$\vec{w} \cdot \vec{v}_1 = 0, \vec{w} \cdot \vec{v}_2 = 0, \ldots, \vec{w} \cdot \vec{v}_k = 0.$$

But this is just a system of k homogeneous linear equations in n variables, with $k \leq n$.

Examples

Find the orthogonal complement of:

(a) $V = \{(t, -2t, 3t) \mid t \in \mathbb{R}\} \subseteq \mathbb{R}^3$.

(b) $V = \{(x, y, z) \mid 2x - y + 3z = 0\} \subseteq \mathbb{R}^3$.