MATHEMATICS FOR SCIENCE STUDENTS

An open-source book

Written, illustrated and typeset (mostly) by

PELEG BAR SAPIR

with contributions from others

$$a^b = e^{b\log(a)}$$

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

$$T\left(\alpha\vec{u} + \beta\vec{v}\right) = \alpha\,T\left(\vec{u}\right) + \beta\,T\left(\vec{v}\right)$$

$$A = Q\Lambda Q^{-1}$$

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\,\mathrm{d}t$$

$$R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

$$\langle \hat{e}_i, \hat{e}_j \rangle = \delta_{ij}$$

$$\vec{v} = \sum_{i=1}^{n} \alpha_i \hat{e}_i$$

$$e^{\pi i} + 1 = 0$$

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)$$

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$$

PUBLISHED IN THE WILD

> **❶ To write/do ❶**
>
> Rights, lefts, etc. will be written here in the future

HERE BE TABLE

CHAPTER
0

INTRODUCTION

In this chapter we introduce key concepts that will be used in later chapters. For this reason, unlike other chapters it contains many statements, sometimes given without thorough explanations or reasoning. While all of these statements are grounded in deep ideas and can be formulated in a rigorous manner, it is advised to first get an intuitive understanding of the ideas before diving into their more formal construction.

> **Note 0.1 In case you are already familiar with the topics**
>
> It is recommended for readers who are familiar with the topics to at least gloss over this chapter and make sure they know and understand all the concepts presented here. **!**

CHAPTER

1

LINEAR ALGEBRA
(INTUITIVE APPROACH)

Linear algebra is one of the most important and often used fields, both in theoretical and applied mathematics. It brings together the analysis of systems of linear equations and the analysis of linear functions (in this context usually called linear transformations), and is employed extensively in almost any modern mathematical field, e.g. approximation theory, vector analysis, signal analysis, error correction, 3-dimensional computer graphics and many, many more.

In this book, we divide our discussion of linear algebra into to chapters: the first (this chapter) deals with a wider, birds-eye view of the topic: it aims to give an intuitive understanding of the major ideas of the topic. For this reason, in this chapter we limit ourselves almost exclusively to discussing linear algebra using 2- and 3-dimensional analysis (and higher dimensions when relevant) using real numbers only. This allows us to first create an intuitive picture of what is linear algebra all about, and how to use correctly the tools it provides us with.

The next chapter takes the opposite approach: it builds all concepts from the ground-up, defining precisely (almost) all basic concepts and proving them rigorously, and only

then using them to build the next steps. This approach has two major advantages: it guarantees that what we build has firm foundations and does not fall apart at any future point, and it also allows us to generalize the ideas constructed during the process to such extent that they can be used as foundation to build ever newer tools we can apply in a wide range of cases.
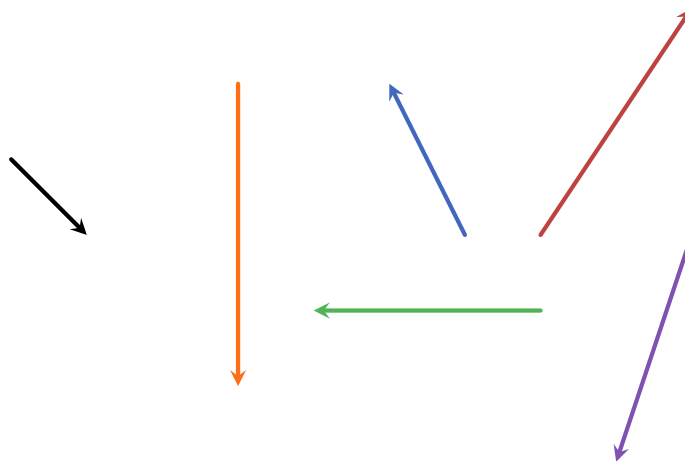
1.1 VECTORS

1.1.1 Basics

Vectors are the fundamental objects of linear algebra: the entire field revolves around manipulation of vectors. In this chapter we deal with the so-called **real vectors**, which can be be defined in a geometric way:
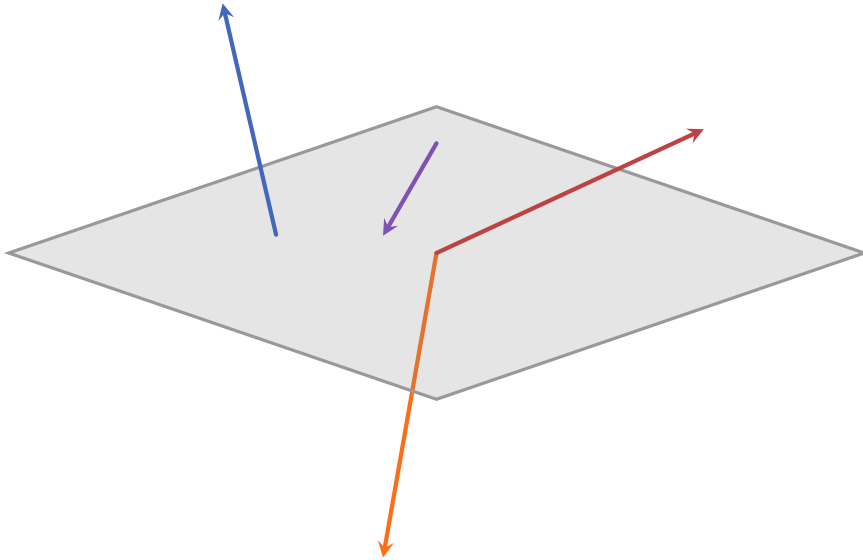
> **Definition 1.1 Real vectors**
>
> A *real vector* is an object with a **magnitude** (also called **norm**) and a **direction**.
>
> π

In this chapter we refer to real vectors simply as *vectors*. Vectors can have $1, 2, 3, \ldots$ number of dimensions. 2-dimensional vectors can be drawn as arrows on a plane:

Similarly, 3-dimensional vectors can be drawn as arrows in space:

Unfortunately, it is difficult (if not impossible) to draw higher-dimensional vectors. For now, we will concentrate on 2-dimensional vectors and explore their properties. Later in the section we will apply what we learned from 2-dimensional vectors to 3-dimensional vectors and higher-dimensional vectors as well.

Vectors are usually denoted in one of the following ways:

- **Arrow above letter**: $\vec{u}$, $\vec{v}$, $\vec{x}$, $\vec{a}$, …

- **Bold letter**: $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{x}$, $\boldsymbol{a}$, …

- **Bar below letter**: $\underline{u}$, $\underline{v}$, $\underline{x}$, $\underline{a}$, …

In this book we use the first notation style, i.e. an arrow above the letter. In addition vectors will almost always be denoted using lowercase *Latin* script.

When discussing vectors in a single context, we always consider them starting at the same point, called the **origin**, and **translating** (moving) vectors around in space does not change their properties: only their norms and directions matter. For example, we can draw the 2-dimensional vectors from before such that their origins all lie on the same point:
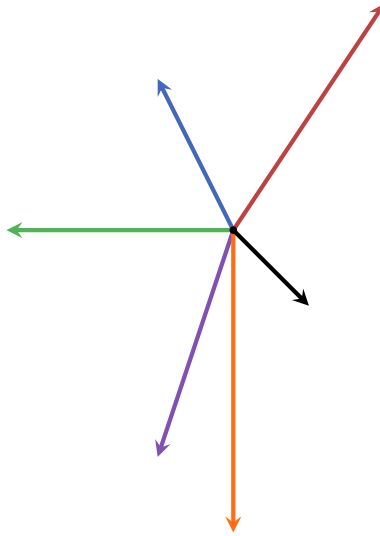
Any vector can be scaled by a real number α: when this happens, its norm is multiplied by α while its direction stays the same. We call α a **scalar**. For example, the vector $\vec{v}$ on the left is scaled here by different scalars $\alpha = 2, 2.5, -1$ and -2:

Note 1.1 Negative scale

As can be seen in the example above, when scaling a vector by a negative amount its direction reverses. However, we consider two opposing direction (i.e. directions that are 180° apart) as being the same direction.

!

In this chapter we use the following notation for the norm of a vector $\vec{v}$: $\left\| \vec{v} \right\|$.

A vector $\vec{v}$ with norm $\left\| \vec{v} \right\| = 1$ is called a **unit vector**, and is usually denoted by replacing the arrow symbol by a hat symbol: $\hat{v}$. Any vector (except $\vec{0}$) can be scaled into a unit vector by scaling the vector by 1 over its own norm, i.e.

$$\hat{v} = \frac{1}{\left\| \vec{v} \right\|} \vec{v}. \tag{1.1.1}$$

The result of normalization is a vector of unit norm which points in the same direction of the original vector.

Two vectors can be added together to yield a third vector: $\vec{u} + \vec{v} = \vec{w}$. To find $\vec{w}$ we use the following procedure (depicted in Figure 1.1):

1. Move (translate) $\vec{v}$ such that its origin lies on the head of $\vec{u}$.

2. The vector $\vec{w}$ is the vector drawn from the origin of $\vec{u}$ to the head of $\vec{v}$.

1 The vectors $\vec{u}$ and $\vec{v}$.

2 Translating $\vec{v}$ such that its origin lies at the head of $\vec{u}$.

3 Drawing the vector $\vec{w}$ from the origin to the head of $\vec{v}$.
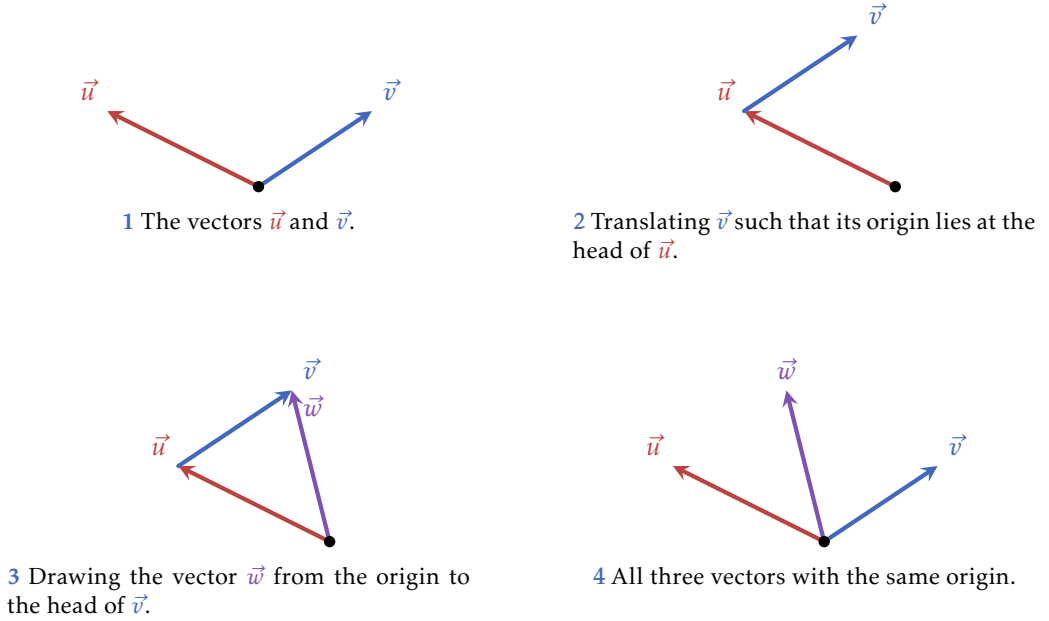
4 All three vectors with the same origin.

Figure 1.1 Vector addition.

The addition of vectors as depicted here is commutative, i.e. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$. This can be seen by using the **parallelogram law of vector addition** as depicted in Figure 1.2: drawing the two vectors $\vec{u}, \vec{v}$ and their translated copies (each such that its origin lies on the other vector's head) results in a parallelogram.

Figure 1.2 The parallelogram law of vector addition.

An important vector is the **zero-vector**, denoted as $\vec{0}$. The zero-vector has a unique property: it is neutral in respect to vector addition, i.e. for any vector $\vec{v}$,

$$\vec{v} + \vec{0} = \vec{v}. \tag{1.1.2}$$

(we also say that $\vec{0}$ is the **additive identity** in respect to vectors.)

Any vector $\vec{v}$ always has an **opposite** vector, denoted $-\vec{v}$. The addition of a vector and its opposite always result in the zero-vector, i.e.

$$\vec{v} + \left(-\vec{v}\right) = \vec{0}. \tag{1.1.3}$$

1.1.2   Column representation

In order to be able to use vectors for actual calculations we must somehow quantify them. When quantifying the length of geometric objects we often first define some unit of measurement, for example a centimeter (cm). We then use this unit to measure the length of different objects.

While this simple approach works well for describing lengths, in the case of vectors we also want to quantify directions - which becomes a bit complicated in higher dimensions if we use angles. Instead, we use more than one unit of measurement; in fact, we use several of them: a vector for each dimension (so in 2-dimensions we use two vectors, in 3-dimensions we use three vectors, etc.). We call these vectors **basis vectors** and collect them in a set we call the **basis set**.

Any two vectors can form a basis set in 2-dimensions **as long as they are not pointing in the same direction**, meaning that they are not scales of eachother: if the vectors are $\vec{u}$ and $\vec{v}$, then there is no $\alpha \in \mathbb{R}$ such that $\vec{u} = \alpha\vec{v}$ (vectors that are scales of eachother are said to be **coliniear**). For example, the following two vectors **can't form a basis in 2-dimensions** since they are coliniear - they point in the same direction (specifically, $\vec{r_2} = 2\vec{r_1}$):
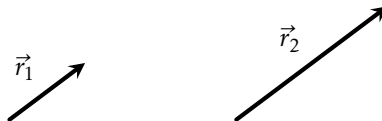
The restriction on coliniearity also means that the zero vector can never be a basis vector, since $\vec{0} = 0 \cdot \vec{v}$ for any $\vec{v}$.
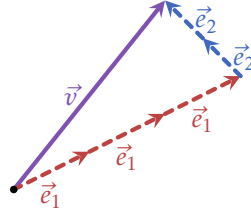
An example of two vectors that form a basis set in 2-dimensions are the following two 2-dimensional vectors $\vec{e_1}$ and $\vec{e_2}$:

> ❗ To write/do ❗
>
> basis vectors in $\mathbb{R}^2$ can't be in the same line

We can then use these basis vectors to measure other 2-dimensional vectors, for example the following vector $\vec{v}$:

We see that using these basis vectors, $\vec{v} = 3\vec{e}_1 + 2\vec{e}_2$. This means that we need to add three times $\vec{e}_1$ and two times $\vec{e}_2$ to construct $\vec{v}$. We denote this fact by writing $\vec{v}$ as a column of two numbers:

$$\vec{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \quad \begin{array}{l} \longleftarrow \text{How much of } \vec{e}_1 \text{ is in } \vec{v} \\ \longleftarrow \text{How much of } \vec{e}_2 \text{ is in } \vec{v} \end{array}$$

Example 1.1 Vectors from basis vectors

Two more vectors represented as sums of the basis vectors $\vec{e}_1$ and $\vec{e}_2$:

$$\vec{u} = 2\vec{e}_1 - 2\vec{e}_2 \quad \longrightarrow \quad \vec{u} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

$$\vec{w} = -3\vec{e}_1 + 1.5\vec{e}_2 \quad \longrightarrow \quad \vec{w} = \begin{bmatrix} -3 \\ 1.5 \end{bmatrix}$$

This notation is frequently refered to as a **column vector**, and the numbers are called the **components** of the vectors. The components themselves are usually denoted using the same symbol used for the vector (without the arrow sign) and an index: for example, the two components of the vector $\vec{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ are $v^1 = 3$ and $v^2 = 2$.

Note 1.2 Index notation

In this book we use upper-index notation (superscript) to denote components of column vectors. Do not mistake these for powers! The reason for this choice (as opposed to lowe-index/subscript notation, i.e. v_1, v_2, etc.) is to stay consistent with later parts of the book, where covectors and in general tensors are presented.

It should now be clear why we restrict basis vectors to be noncolinear: if we choose two coliniear vectors, any addition of the two vectors will end up pointing in the same direction as them, and we won't be able to express all vectors as columns (we will be restricted to only a single direction).

> **Note 1.3 Coordinate systems**
>
> The set of basis vectors used to represent vectors as columns is sometimes called a **coordinate system**. We will use the name **basis set** to stay consistent. !

Vectors have different components when using different basis sets. For example, we can use the following two vectors $\vec{d}_1$ and $\vec{d}_2$ as basis vectors:

In this new basis set, the vector $\vec{v}$ from earlier has the following column representation:

$$\vec{v} = 2\vec{d}_1 + 2.5\vec{d}_2 \quad \longrightarrow \quad \vec{v} = \begin{bmatrix} 2 \\ 2.5 \end{bmatrix}$$
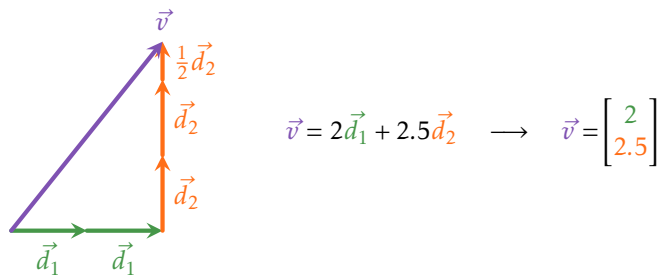
(i.e. its components are $v^1 = 2$ and $v^2 = 2.5$)

This brings us to an important idea, which unfortunately might confuse those who are new to the topic: **vectors and their column representation are two separate things!** Vectors are objects with a length (norm) and a direction. They don't "care" about how *we* describe them numerically: no matter what basis set we use, vectors remain the same - it is their *representation* which changes.

In fact, not only does the choice of basis set affect the column representation of all vectors[1], two different vectors **can have the same column representation using two different basis sets..** For example. let's say we choose any two non-zero vectors to be used as basis vectors: $\vec{u}$ and $\vec{v}$. Then the following is always true:

$$\vec{u} = 1 \cdot \vec{u} + 0 \cdot \vec{v},$$
$$\vec{v} = 0 \cdot \vec{u} + 1 \cdot \vec{v}.$$

---

[1] except the zero vector, which is always $\vec{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Therefore, the column representations of $\vec{u}$ and $\vec{v}$ will be exactly $\vec{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. This means that the column vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ can represent **any** two (non-zero) vectors we wish! So remember: you must always be sure with which basis vectors you are working. Otherwise, mistakes are bound to happen and calculations might make no sense.

> ❶ To write/do ❶
>
> more examples?

Since we need two real numbers to express any such 2-dimensional vector as a column, we call the set of all such vectors $\mathbb{R} \times \mathbb{R}$, or simply $\mathbb{R}^2$.

### 1.1.3   Vector operations and the zero vector in column form

Previously we saw how to scale and add vectors (the latter using the parallogram method). Let us now see how we perform these operations using the column representation of vectors. We will use the same basis vectors as before:
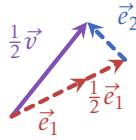

We start with the vector $\vec{v}$ from before, i.e. $3 \cdot \vec{v} = \vec{e_1} + 2 \cdot \vec{e_2}$ and $\vec{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$:
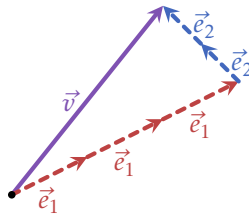
What then would be the components of, say, $\frac{1}{2}\vec{v}$? First we scale $\vec{v}$ by $\frac{1}{2}$: remember, this just means "squeezing" the vector to $\frac{1}{2}$ its former length, keeping it pointing in the same direction:
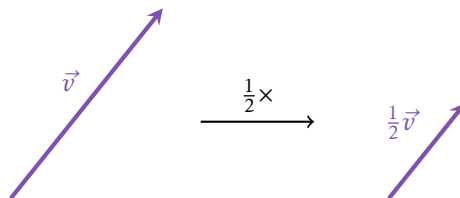
Now we use the basis vectors $\vec{e}_1$ and $\vec{e}_2$ to get the column representation of $\frac{1}{2}\vec{v}$:


We get that $\frac{1}{2}\vec{v} = 1.5\vec{e}_1 + 1\vec{e}_2$, i.e. $\frac{1}{2}\vec{v} = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \cdot 3 \\ \frac{1}{2} \cdot 2 \end{bmatrix}$ - i.e. scaling $\vec{v}$ by $\frac{1}{2}$ simply multiplied both its components by $\frac{1}{2}$.

This is in fact true for any vector and any scalar, using any basis set: scaling a vector by a scalar α results in multiplying each of its components by α:

$$\vec{u} = \begin{bmatrix} u^1 \\ u^2 \end{bmatrix} \quad \Longrightarrow \quad \alpha\vec{u} = \begin{bmatrix} \alpha \cdot u^1 \\ \alpha \cdot u^2 \end{bmatrix}. \tag{1.1.4}$$

Now let's add two vectors together: $\vec{v} - \vec{w}$ (the same $\vec{w}$ from Example 1.1), again using the same basis vectors. We saw that $\vec{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} -3 \\ 1.5 \end{bmatrix}$. Using the parallelogram method their addition is the following vector $\vec{a}$:
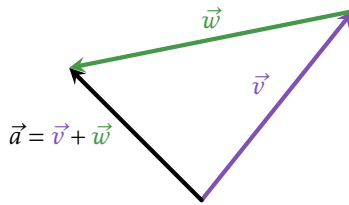
Now we calculate the components of $\vec{a}$ in the basis $\vec{e}_1, \vec{e}_2$:
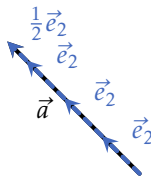
We see that $\vec{a} = 0 \cdot \vec{e}_1 + 3.5\vec{e}_2 = \begin{bmatrix} 0 \\ 3.5 \end{bmatrix} = \begin{bmatrix} 3 + (-3) \\ 2 + 1.5 \end{bmatrix}$, so the column representation of the addition of two vectors is simply the addition of their components.

This is infact not a coincidence: given two generic column vectors in the same basis $\{\vec{e}_1, \vec{e}_2\}$,

$$\vec{u} = \begin{bmatrix} u^1 \\ u^2 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} v^1 \\ v^2 \end{bmatrix},$$

we can write them as follows:

$$\vec{u} = u^1\vec{e}_1 + u^2\vec{e}_2,$$
$$\vec{v} = v^1\vec{e}_1 + v^2\vec{e}_2.$$

When adding them in this form we can collect all similar terms together:

$$\vec{w} = \vec{u} + \vec{v} = u^1\vec{e}_1 + u^2\vec{e}_2 + v^1\vec{e}_1 + v^2\vec{e}_2$$
$$= \left(u^1 + v^1\right)\vec{e}_1 + \left(u^2 + v^2\right)\vec{e}_2.$$

and the column of the resulting vector $\vec{w}$ is therefore

$$\vec{w} = \begin{bmatrix} u^1 + v^1 \\ u^2 + v^2 \end{bmatrix},$$

which is exactly the **column-wise** addition of $\vec{u}$ and $\vec{v}$.

Note 1.4 Scaling and addition using different basis sets

This is only true if the column representation of the two vectors uses the same basis set. Adding two column vectors in different basis sets requires changing one of the column representation to the basis of the other one. We will deal with change of basis later in the chapter.

!

Negating a vector can be thought of as multiplying it by the scalar $\alpha = -1$. Therefore, given a vector $\vec{v} = \begin{bmatrix} v^1 \\ v^2 \end{bmatrix}$, its opposite would be

$$-\vec{v} = \begin{bmatrix} -v^1 \\ -v^2 \end{bmatrix}. \tag{1.1.5}$$

And finally, since $\vec{0} = \vec{v} + \left(-\vec{v}\right)$ for any $\vec{v}$, we get that

$$\vec{0} = \begin{bmatrix} v^1 \\ v^2 \end{bmatrix} + \left(-1\begin{bmatrix} v^1 \\ v^2 \end{bmatrix}\right) = \begin{bmatrix} v^1 \\ v^2 \end{bmatrix} + \begin{bmatrix} -v^1 \\ -v^2 \end{bmatrix} = \begin{bmatrix} v^1 - v^1 \\ v^2 - v^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{1.1.6}$$

Summary 1.1 Column representation of vectors

Scaling and adding vectors using their column representation is rather simple:

1. Scaling a vector by a scalar α is done by multiplying each of the vector's components by α: $\alpha\begin{bmatrix} v^1 \\ v^2 \end{bmatrix} = \begin{bmatrix} \alpha v^1 \\ \alpha v^2 \end{bmatrix}$.

2. Adding two vectors is done by adding their respective components: $\begin{bmatrix} u^1 \\ u^2 \end{bmatrix} +$


$\begin{bmatrix} v^1 \\ v^2 \end{bmatrix} = \begin{bmatrix} u^1 + v^1 \\ u^2 + v^2 \end{bmatrix}$. In general, we say that using the column representation, vector scaling and addition is done **component-wise**.

3. Using the above operations we get that the opposite of a vector $\begin{bmatrix} v^1 \\ v^2 \end{bmatrix}$ is $\begin{bmatrix} -v^1 \\ -v^2 \end{bmatrix}$.

4. We also get that the zero vector is always $\vec{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

1.1.4 Cartesian coordinates and the standard basis set

We can place vectors on a two-dimensional Cartesian coordinate system, such that their origin coincide with the axis-origin (the point $O = (0,0)$). We then mark the point where its head lies as $p = (p_x, p_y)$:
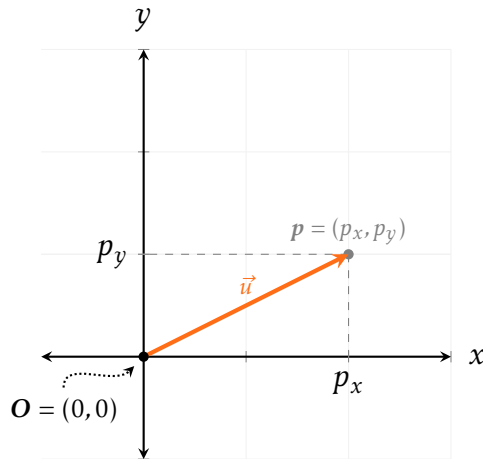
This arrangement has a fitting basis set, which we call the **standard basis set of $\mathbb{R}^2$**:

1. $\hat{x}$: a vector of unit length in the direction of the x-axis, i.e. from the origin of the axes to point $(1, 0)$.

2. $\hat{y}$: a vector of unit length in the direction of the y-axis, i.e. from the origin of the axes to point $(0, 1)$.
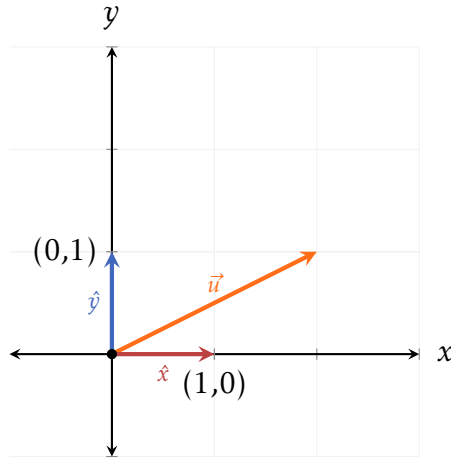
Using the standard basis set, the first component of any vector is simply p_x and the second component is p_y, i.e. $\vec{u} = \begin{bmatrix} p_x \\ p_y \end{bmatrix}$. For example, we see that the vector $\vec{u}$ in the above figure has the column representation $\vec{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

---

Example 1.2 Vectors on the Cartesian plane

Some more vectors on the 2-dimensions Cartesian plane and their column representation using the standard basis (each grid line is one unit in size):

---
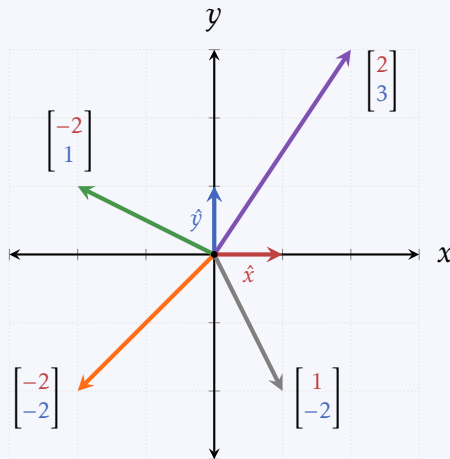
it is common to call u^1 simply x, and u^y simply y, and therefore the components of some vector $\vec{u}$ are called u^x and u^y, respectively.

1.1.5 Polar coordinates

Using the 2-dimensional Cartesian plane we can define an alternative coordinate system for vectors: notice that each vector $\vec{u} = \begin{bmatrix} u^x \\ u^y \end{bmatrix}$ defines a right triangle together with the x-axis; The side on the x-axis and the side parallel to the y-axis are then u^x and u^y, respectively:
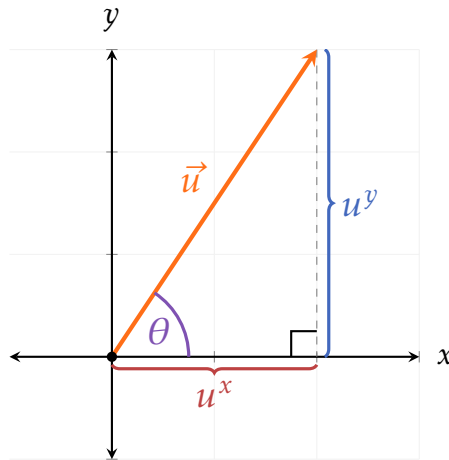
Due to the Pythagorean theorem we know that the norm of $\vec{u}$ is

$$\left\| \vec{u} \right\| = \sqrt{(u^x)^2 + (u^y)^2}, \tag{1.1.7}$$

and due to trigonometry, the angle θ between $\vec{u}$ and the x-axis is

$$\theta = \arctan\left(\frac{u^y}{u^x}\right). \tag{1.1.8}$$

The alternative coordinate system, called the **polar coordinate system**, uses a tuple of numbers: (R, θ) - where $R = \left\| \vec{u} \right\|$. The conversion from the coordinates (R, θ) to the column form $\begin{bmatrix} u^x \\ u^y \end{bmatrix}$ is therefore

$$u^x = R\cos(\theta),$$
$$u^y = R\sin(\theta).$$

(cf. **??**)

> ❗ To write/do ❗
>
> examples of transitioning between coords sys

1.1.6 Unit vectors in the standard basis

> **❶ To write/do ❶**
>
> this subsection

1.1.7 Examples

Before we continue with the topic, let's use what we learned to solve some simple problems:
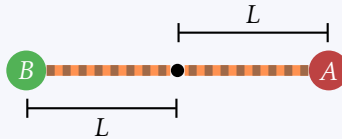
> **Example 1.3 Turtle walk**
>
> Joe the turtle went on a stroll in the field: he first walked 10 meters, then turned 30° counter clockwise and walked 5 more meters. After a short rest Joe noticed the pond to his right, so he turned 45° clockwise and walked to the pond - about 30 meters away. When he reached the pond he swam across it at a 20° counter clockwise and 10 meters forward.

> **Example 1.4 Force on an onject**
>
> Consider a simplified game of "tug of war", where two people A and B are pulling on some object (black circle) using two ropes of the same length L directly opposite to one another[a]:
>
>
>
> If both people pull with the exact same force, the object will stay at the same place. We can explain this using vector addition: if we represent the force applied on the object by the pull of person A using the vector $\vec{F}$, then the pull force applied on the object by person B is $-\vec{F}$, since the total force must add up to zero (no force=no change in position[b]).
>
> Now consider adding a third person pulling on the object, and that all three people are arranged such that they are at 120° to eachother and at the exact same length L from the object they are pulling:
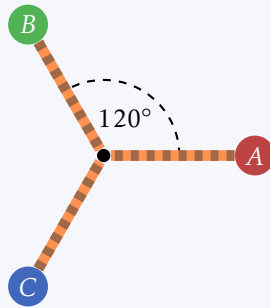
If the three people are pulling with the same force, what would be the total force experienced by the object? Again, we can express the forces using vectors: $\vec{F}_A$, $\vec{F}_B$ and $\vec{F}_C$ - we want to calculate the sum $\vec{F}_A + \vec{F}_B + \vec{F}_C$. Since all forces have the same magnitude, we'll call that magnitude simply F. Using the standard basis vectors in $\mathbb{R}^2$ we can consider the vector $\vec{F}_A$ to be pointing to the right and thus having the column representation

$$\vec{F}_A = \begin{bmatrix} F \\ 0 \end{bmatrix}.$$

The vector $\vec{F}_B$ has the same magnitude but a different direction: we can use polar coordinates to find its column representations, and we get that these are $(F, 120°)$. A simple conversion yields

$$\vec{F}_B = \begin{bmatrix} F\cos(120°) \\ F\sin(120°) \end{bmatrix} = F \begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix} = \frac{F}{2} \begin{bmatrix} -1 \\ \sqrt{3} \end{bmatrix}.$$

Similar calculation for $\vec{F}_C$ gives $\vec{F}_C = \frac{F}{2} \begin{bmatrix} -1 \\ -\sqrt{3} \end{bmatrix}$. Adding all three vectors together gives

$$\vec{F}_A + \vec{F}_B + \vec{F}_C = \begin{bmatrix} F \\ 0 \end{bmatrix} + \frac{F}{2} \begin{bmatrix} -1 \\ \sqrt{3} \end{bmatrix} + \frac{F}{2} \begin{bmatrix} -1 \\ -\sqrt{3} \end{bmatrix}$$
$$= \begin{bmatrix} F - \frac{F}{2} + \frac{F}{2} \\ 0 + \frac{\sqrt{3}F}{2} - \frac{\sqrt{3}F}{2} \end{bmatrix}$$
$$= \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

We get, once again, that the total force applied on the object is zero.
At this point one might wonder if this pattern always holds, that is - given $n \geq 1$ people arranged at the vertices of a regular polygon, each pulling with the same force on a rope connected to some object at the center of the regular polygon, will the total force be zero? To help picture this question, here are schematics for the cases $n = 4, 5, 6, 7$:

$\underline{n = 4}$

$\underline{n = 5}$

90°

72°

$\underline{n = 6}$

$\underline{n = 7}$

60°

51°
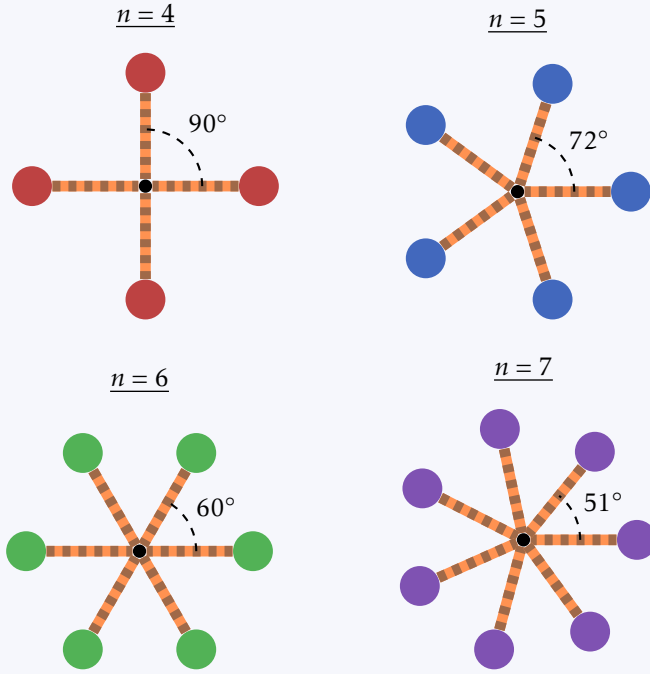
> ❗ To write/do ❗
>
> correct 51° issue

Intuitively, one would expect that for any $n \geq 1$, $\sum\limits_{i=1}^{n} \vec{F}_i = \vec{0}$ due to symmetry. Let's do the actual "dirty work" and prove this. First, the angle between each two consecutive ropes/people is $\frac{360°}{n}$. We will use $\frac{2\pi}{n}$ instead, because radians are simply *better* than degrees. We get that each vector $\vec{F}_k$ has the following polar coordinates:

$$\vec{F}_k : \left(L, \frac{2\pi}{n}k \right),$$

where $k = 0, 1, 2, \ldots, n-1$ - we use this so that the first vector always points directly to the right, which makes the calculations a bit more convinent.

Using these polar coordinates, we can write each vector in column form in the standard basis in $\mathbb{R}^2$:

$$\vec{F}_k = \begin{bmatrix} L\cos\left(\frac{2\pi}{n}k\right) \\ L\sin\left(\frac{2\pi}{n}k\right) \end{bmatrix}.$$

Therefore, the sum of all n vectors is

$$\vec{V} = \sum_{k=0}^{n-1} F_k = \sum_{k=0}^{n-1} \begin{bmatrix} L\cos\left(\frac{2\pi}{n}k\right) \\ L\sin\left(\frac{2\pi}{n}k\right) \end{bmatrix} = L \sum_{k=0}^{n-1} \begin{bmatrix} \cos\left(\frac{2\pi}{n}k\right) \\ \sin\left(\frac{2\pi}{n}k\right) \end{bmatrix}.$$

Since we want to show that the sum is zero, we can ignore the norm L as it doesn't play a role.

Since column vectors are added *component-wise*, we actually need to show the following two equalities:

$$1. \quad \sum_{k=0}^{n-1} \cos\left(\frac{2\pi}{n}k\right) = 0,$$

$$2. \quad \sum_{k=0}^{n-1} \sin\left(\frac{2\pi}{n}k\right) = 0.$$

Starting with the first sum, we will use Lagrange's trigonometric identities (**??**). Since these identities involve sums with $k = 0, 1, 2, \ldots, n$ but our sums only use $k = 0, 1, 2, \ldots, n-1$, we would need to subtract the term where $k = n$ from the result, that is the term $\cos\left(\frac{2\pi}{n}n\right) = \cos(2\pi) = 1$. Using $\theta = \frac{2\pi}{n}$ in the identities we get that

$$\sum_{k=0}^{n-1} \cos\left(\frac{2\pi}{n}k\right) = \sum_{k=0}^{n} \cos\left(\frac{2\pi}{n}k\right) - 1$$

$$= \frac{\sin\left(\frac{1}{2}\frac{2\pi}{n}k\right) + \sin\left(\left[n+\frac{1}{2}\right]\frac{2\pi}{n}k\right)}{2\sin\left(\frac{1}{2}\frac{2\pi}{n}k\right)} - 1$$

$$= \frac{\sin\left(\frac{\pi}{n}k\right) + \sin\left(2\pi k + \frac{\pi}{n}k\right)}{2\sin\left(\frac{\pi}{n}k\right)} - 1.$$

We can use the angle sum trigonometric identities (**??**) to calculate $\sin\left(2\pi k + \frac{\pi}{n}k\right)$:

$$\frac{\sin\left(\frac{\pi}{n}k\right) + \sin\left(2\pi k + \frac{\pi}{n}k\right)}{2\sin\left(\frac{\pi}{n}k\right)} - 1 = \frac{\sin\left(\frac{\pi}{n}k\right) + \left[\sin(2\pi k)\cos\left(\frac{\pi}{n}k\right) + \cos(2\pi k)\sin\left(\frac{\pi}{n}k\right)\right]}{2\sin\left(\frac{\pi}{n}k\right)} - 1$$

$$= \frac{\sin\left(\frac{\pi}{n}k\right) + \sin\left(\frac{\pi}{n}k\right)}{2\sin\left(\frac{\pi}{n}k\right)} - 1$$

$$= \frac{2\sin\left(\frac{\pi}{n}k\right)}{2\sin\left(\frac{\pi}{n}k\right)} - 1$$

$$= 1 - 1$$

$$= 0.$$

And for the second sum (using $\sin\left(\frac{2\pi}{n}n\right) = \sin(2\pi) = 0$):

$$\sum_{k=0}^{n-1} \sin\left(\frac{2\pi}{n}k\right) = \sum_{k=0}^{n} \sin\left(\frac{2\pi}{n}k\right)$$

$$= \frac{\cos\left(\frac{1}{2}\frac{2\pi}{n}k\right) - \cos\left(\left[n+\frac{1}{2}\right]\frac{2\pi}{n}k\right)}{2\sin\left(\frac{1}{2}\frac{2\pi}{n}k\right)}$$

$$= \frac{\cos\left(\frac{\pi}{n}k\right) - \cos\left(2\pi k + \frac{\pi}{n}k\right)}{2\sin\left(\frac{\pi}{n}k\right)}$$

$$= \frac{\cos\left(\frac{\pi}{n}k\right) - \left[\cos(2\pi k)\cos\left(\frac{\pi}{n}k\right) - \sin(2\pi k)\sin\left(\frac{\pi}{n}k\right)\right]}{2\sin\left(\frac{\pi}{n}k\right)}$$

$$= \frac{\cos\left(\frac{\pi}{n}k\right) - \cos\left(\frac{\pi}{n}k\right)}{2\sin\left(\frac{\pi}{n}k\right)}$$

$$= 0.$$

Indeed, adding up all the vectors results in the zero vector - i.e. no force is applied on the object.

---

[a]A key to solving all mathematical problems is being able to draw a simplified schematic of the situation. That is most definitely *not* an excuse for the lousy drawing…

[b]Yes, technically no force means no acceleration, so if this bothers you just assume the object had no initial velocity.

1.1.8   3-dimensional vectors and the space $\mathbb{R}^3$

Everything we discussed so far regarding vectors in $\mathbb{R}^2$ can be easily applied to vectors existing in 3-dimensional space, which we also call $\mathbb{R}^3$ (i.e. $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$). These vectors also have a length (norm) and a direction, and can be scaled and added in the exact same way as vectors in $\mathbb{R}^2$.

The main difference between vectors in $\mathbb{R}^2$ and vectors in $\mathbb{R}^3$ is that any basis set in $\mathbb{R}^3$ must contain **three** basis vectors instead of two, and these must not lie in the same *plane* (we will explore exactly what that means a bit later). The column representation of vectors in $\mathbb{R}^3$ therefore has three entries. For example, given the three basis vectors $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$, then

$$\vec{v} = \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$$

is the column representation of the vector

$$\vec{v} = 2\vec{e}_1 - 3\vec{e}_2 + 5\vec{e}_3.$$

The restriction on the set of vectors which can be used as basis vectors in $\mathbb{R}^3$ is that **they must not be on the same plane**, as that will restrict the columns they can represent to also be on the same plane instead of pointing in all possible direction in 3-dimensions. We call such vectors coplanar. Note that if two or more vectors are on the same line in 3-dimensions, they are also on the same plane. This means that any two of the three basis vectors (or all of them) are also not allowed to be coliniear.
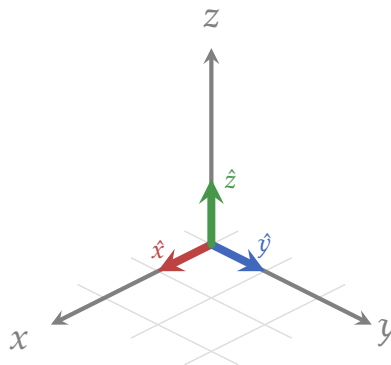
> ❗ To write/do ❗
>
> figure of coplanar vectors?

Addition of column vectors in $\mathbb{R}^3$ is also done component-wise, for example:

$$\begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix} + \begin{bmatrix} 0 \\ 4 \\ -2 \end{bmatrix} = \begin{bmatrix} 2+0 \\ -3+4 \\ 5-2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}.$$

$\mathbb{R}^3$ also has a standard basis set: using a 3-dimensional Cartesian coordinate system with the axes x, y and z,

When using these basis vectors the three components of a vector $\vec{u}$ are sometimes refered to as u^x, u^y and u^z, and its norm is

$$\|\vec{u}\| = \sqrt{(u^x)^2 + (u^y)^2 + (u^z)^2}. \tag{1.1.9}$$

Equation 1.1.9 is a direct result of the Pythagorean theorem:

> **Proof 1.1 Pythagorean theorem in 3D**
>
> Given a box with sides a, b and c, the length of the diagonal line of the base rectangle (in the figure below d) is given by the Pythagorean theorem (see figure for reference):
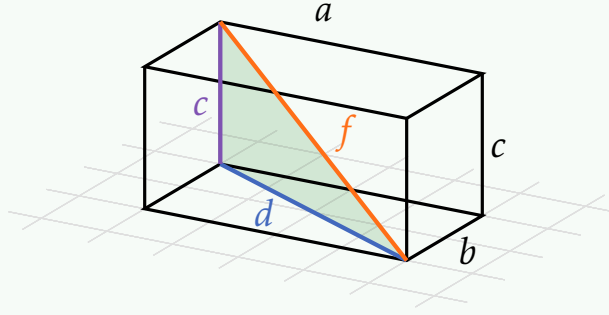
$$d = \sqrt{a^2 + b^2}.$$

Using the height of the box c we can easily calculate the length of the main diagonal of the box, f, using the Pythagorean theorem on the right triangle with sides d, c and hypothenous f (shaded in green in the figure):

$$f = \sqrt{d^2 + c^2} = \sqrt{\left(\sqrt{a^2 + b^2}\right)^2 + c^2} = \sqrt{a^2 + b^2 + c^2}.$$

QED

Using the above proof we simply make the following substitutions:

$$u^x = a,$$
$$u^y = b,$$
$$u^z = c.$$

The norm of $\vec{u}$ is then $\left\| \vec{u} \right\| = f$.

> **Note 1.5   Yet another notation** ¯_(ツ)_/¯
>
> For whatever reason, the notation $\hat{i}$, $\hat{j}$ and $\hat{k}$ is sometimes used in physics to refer to the standard basis vectors. Accordingly, the components of a vector $\vec{u}$ are refered to as u^i, u^j and u^k.
>
> !

1.1.9 n-dimensional vectors and the space $\mathbb{R}^n$

Of course, why stop at three dimensions? We can generalize everything which worked so far to any n-dimensional space where $n = 4, 5, 6, \ldots$ - i.e. the space $\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R} = \mathbb{R}^n$. Picturing vectors in 4-dimensions and above is rather difficult, but the logic remain the same: vectors still have a length and a direction, they can be scaled by a real number and added together in exactly the same fashion as in $\mathbb{R}^2$ and $\mathbb{R}^3$. The zero vector $\vec{0}$ exists and is neutral to addition, etc.

Using the column representation in $\mathbb{R}^n$ makes calculations rather easy: we need to choose n basis vectors and use them to represent any other vector. We can't use any

vectors though - just like basis vectors in $\mathbb{R}^2$ must not be on the same *line*, and basis vectors in $\mathbb{R}^3$ can't be in the same *plane*, so do the basis vectors in $\mathbb{R}^n$ not be in the same $(n-1)$-dimensional *space*. We will explore what that means exactly in the next subsection, but for now let us assume that this is always true, even though it might be somewhat abstract right now.

Like with $\mathbb{R}^2$ and $\mathbb{R}^3$, we will name the n vectors we choose as basis vectors for $\mathbb{R}^n$ using lower-index notation:

$$\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}.$$

Then, we can express any vector in $\mathbb{R}^n$ using these basis vectors. For example, in $\mathbb{R}^7$ the vector

$$\vec{v} = 2\vec{e_1} - 5\vec{e_2} + \vec{e_3} - 2\vec{e_4} - 9\vec{e_5} + \vec{e_6} - 2\vec{e_7}$$

will be written in column representation as

$$\vec{v} = \begin{bmatrix} 2 \\ -5 \\ 1 \\ -2 \\ -9 \\ 1 \\ -2 \end{bmatrix}.$$

Generally of course, the i-component of a vector $\vec{u} \in \mathbb{R}^n$ is denoted as u^i (again, note the upper-index notation and don't confuse it with an exponent/power). So in the case above, $v^3 = 1$ and $v^5 = -9$, for example.

Like before, scaling and adding vectors in column representation in $\mathbb{R}^n$ is done component-wise: given the two vectors

$$\vec{u} = \begin{bmatrix} u^1 \\ u^2 \\ \vdots \\ u^n \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{bmatrix},$$

and scalar $\alpha \in \mathbb{R}$, the following holds:

$$\alpha\vec{u} = \begin{bmatrix} \alpha u^1 \\ \alpha u^2 \\ \vdots \\ \alpha u^n \end{bmatrix}, \tag{1.1.10}$$

$$\vec{u} + \vec{v} = \begin{bmatrix} u^1 + v^1 \\ u^2 + v^2 \\ \vdots \\ u^n + v^n \end{bmatrix}. \tag{1.1.11}$$

The standard basis set in $\mathbb{R}^n$ is a set of vectors each of norm 1 and pointing in the direction of the n-th Cartesian axis. Again, quite difficult to picture for $n \geq 4$, but everything

basically behaves just like in $\mathbb{R}^2$ and $\mathbb{R}^3$. The notation for the standard basis vectors in $\mathbb{R}^n$ that we use in this book is the following:

$$\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_n. \tag{1.1.12}$$

The norm of a vector $\vec{v} \in \mathbb{R}^n$ is a generalization of the Pythagorean theorem:

$$\left\| \vec{v} \right\| = \sqrt{\left(v^1\right)^2 + \left(v^2\right)^2 + \cdots + (v^n)^2} = \sqrt{\sum_{k=1}^{n} \left(v^k\right)^2}. \tag{1.1.13}$$

(just as before, the 2 above the components outside the paranthesis denotes power, and the numbers inside the paranthesis denote index)

1.1.10 Einstein summation notation

We now introduce a very common notation which is used almost everywhere in physics and will give a first hint on why we insist on the upper-index notation for vector components: the **Einstein summation notation**. This notation does the following: everytime we see together an index both as upper-index and as lower-index *in the same term*, it means that there is an implicit sum over that index (the limits of the sum are derived from context).

For example, instead of writing a general vector $\vec{v} \in \mathbb{R}^n$ as

$$\vec{u} = u^1 \vec{e}_1 + u^2 \vec{e}_2 + \cdots + u^n \vec{e}_n = \sum_{k=1}^{n} u^k \vec{e}_k,$$

we simply write

$$\vec{u} = u^k \vec{e}_k. \tag{1.1.14}$$

Another example: using the Einstein summation notation, the norm of the above vector $\vec{u}$ is written simply as

$$\left\| \vec{u} \right\| = \sqrt{u^k \vec{e}_k}. \tag{1.1.15}$$

This notation might seem arbitrary and annoying now, but it becomes incredibly helpful when discussing high-dimensional **tensors** which might have several components of different types (for which we use both upper- and lower-indices).

> ❗ To write/do ❗
>
> examples for Einstein notation

> ❗ To write/do ❗
>
> unit vectors and normalization in column representation

1.1.11 Linear combinations and subspaces

> ❶ To write/do ❶
>
> 1. linear combs
>
> 2. spans
>
> 3. basis sets and dimensionality
>
> 4.

1.1.12 The dot product

> ❶ To write/do ❶
>
> an intro with a daily example of projection

> ❶ To write/do ❶
>
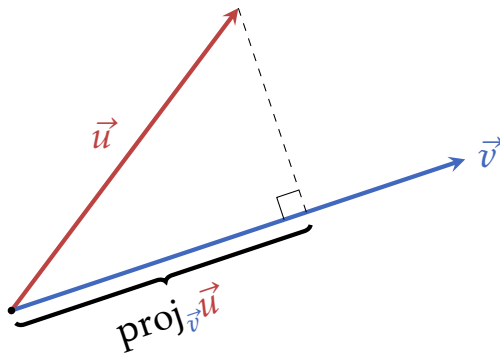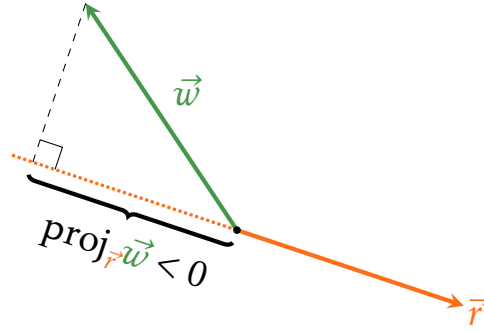> angle between two vectors in $\mathbb{R}^n$?

> ❶ To write/do ❶
>
> don't forget to mention that the dot product calculation only works in orthonormal basis sets!

An **orthogonal projection** of a vector $\vec{u}$ onto another vector $\vec{v}$ essentially measures how much of $\vec{u}$ is in the direction of $\vec{v}$:


Depending on the directions of both vectors, the projection might be negative:

Of course, if the the vectors are **orthogonal** then the projection equals zero, since the projected vector and the projection line coincide. On the other hand, if the two vectors are pointing in the same direction, the projection would equal the norm of the projected vector.

Since the two vectors form a right triangle together with the perpendicular line, we can use the angle θ between them to calculate the projection:
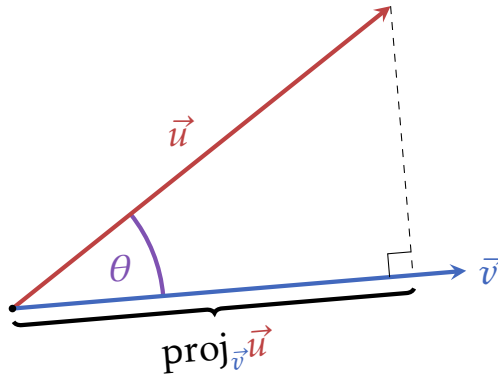
By applying **??** we get the following expression:

$$\text{proj}_{\vec{v}}\vec{u} = \left\| \vec{u} \right\| \cos(\theta). \tag{1.1.16}$$

Using the above equation we can define a product between the two vectors, called the **dot product** - it simply the orthogonal projection of $\vec{u}$ onto $\vec{v}$ multiplied by the norm of $\vec{v}$:

$$\vec{u} \cdot \vec{v} = \left\| \vec{v} \right\| \text{proj}_{\vec{v}}\vec{u} = \left\| \vec{u} \right\| \left\| \vec{v} \right\| \cos(\theta). \tag{1.1.17}$$

> **Note 1.6   The name "dot product"**
>
> Yes, the term "dot product" comes from the use of the symbol $\cdot$ to represent it.
>
> **!**

The dot product works in any space $\mathbb{R}^n$: it takes two vectors and returns a scalar, i.e. it can be thought of as a function of the type

$$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \tag{1.1.18}$$

and therefore, depending mostly on context, it is also called **scalar product**. Another name sometimes used is **inner product**, which is a more generalized concept that we introduce in future chapters.

> **Note 1.7   Dot product notations**
>
> While in this section we use the dot notation, other notations are also used depending on the context. Here are some examples for such notations for two vectors $\vec{a}$ and $\vec{b}$:
>
> | Notation | Remarks |
> |---|---|
> | $\left\langle \vec{a}, \vec{b} \right\rangle$ | Used in both abstract and applied mathematics. |
> | $\left(\vec{a}, \vec{b} \right)$ | Mostly used in abstract mathematics. |
> | $\left\langle \vec{a} \middle\| \vec{b} \right\rangle$ | Called the **Dirac notation**, also known as the **bra-ket notation**. Mostly used in physics, especially in quantum physics. This is actually a very useful notation, and we discuss it more in-depth in a later section. |
>
> **!**

Some properties of the dot product:

• **It is commutative**: since $\vec{u}$, $\vec{v}$ and $\cos(\theta)$ are all real numbers, we can flip the order of the dot product and get the same result:

$$\vec{u} \cdot \vec{v} = \left\| \vec{u} \right\| \left\| \vec{v} \right\| \cos(\theta) = \left\| \vec{v} \right\| \left\| \vec{u} \right\| \cos(\theta) = \vec{v} \cdot \vec{u}.$$

• **It is distributive over vector addition**: $\vec{u} \cdot \left(\vec{v} + \vec{w} \right) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$.

• **We can pull scalars out of the product**:

$$\left(\alpha \vec{u} \right) \cdot \vec{v} = \left\| \alpha \vec{u} \right\| \left\| \vec{v} \right\| \cos(\theta)$$

$$= \alpha \|\vec{u}\| \|\vec{v}\| \cos(\theta) \leftarrow \alpha\left(\vec{u} \cdot \vec{v}\right)$$
$$= \|\vec{u}\| \|\alpha\vec{v}\| \cos(\theta). \leftarrow \vec{u} \cdot \left(\alpha\vec{v}\right)$$

And now for the most important property: as we mentioned before, when the two vectors are orthogonal - the dot product is zero. This can be seen arithmetically using Equation 1.1.17: when the vectors are orthogonal $\theta = 90° = \frac{\pi}{2}$, and thus $\cos(\theta) = 0$. If both vectors are non-zero and their dot product is zero - they are orthogonal.

This is such an important fact that we will put effort into framing it nicely, so you could memorize it well. How well should you memorize this? So well that if someone wakes you up in the middle of the night and asked you, you could easily repeat it!

$$\vec{u} \cdot \vec{v} \;=\; 0$$
$$\text{(and both are non-zero)}$$

$$\Updownarrow$$

$$\vec{u} \text{ and } \vec{v} \text{ are orthogonal}$$
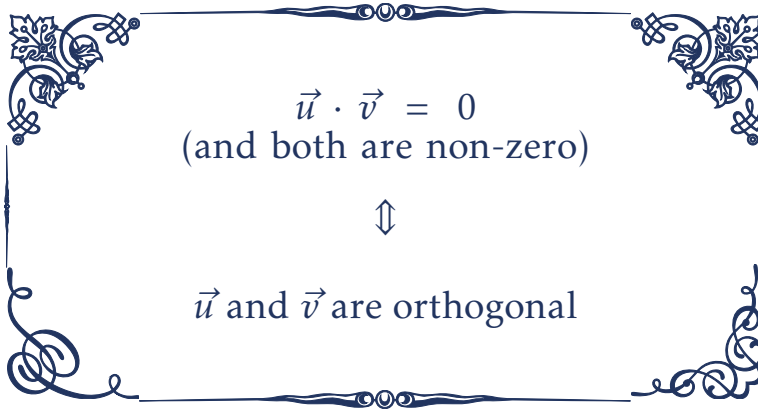
❗ To write/do ❗

calculation examples?

A result of the fact that the dot product of two vectors can be zero even if both vectors are non-zero is that unlike with real numbers, terms can't be cancelled: suppose we have some three **non-zero** real numbers a, b and c, such that

$$ab = ac,$$

then we know that $b = c$. However, that doesn't work with vectors and the dot product: suppose the three vectors $\vec{u}$, $\vec{v}$ and $\vec{w} \neq \vec{v}$ are all non-zero, and

$$\vec{u} \cdot \vec{v} = \vec{u} \cdot \vec{w}.$$

By rearranging and using the distributive property of the dot product, we can write

$$\vec{u} \cdot (\vec{v} - \vec{w}) = \vec{0}.$$

This means that $\vec{u}$ is orthogonal to the vector $\vec{u} - \vec{w}$, which is also non-zero since $\vec{w} \neq \vec{v}$.

Of course it is not always easy to find the angle between two vectors - especially not in spaces with more than 2 dimensions. Since in practive we almost always use vectors in

their column represented, it would be nice to find a formula for the dot product which uses this form. Luckily there is such a formula, and it is actually quite simple! Instead of just showing it, let's find it ourselves[2].

We start with two vectors, $\vec{u}$ and $\vec{v}$. We will assume for simplicity that we are working in $\mathbb{R}^3$ - although that is not strictly necessary, and after we find the formula we will see how to generalize it to $\mathbb{R}^n$. Using the standard basis in $\mathbb{R}^3$ we have

$$\vec{u} = \begin{bmatrix} u^x \\ u^y \\ u^z \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} v^x \\ v^y \\ v^z \end{bmatrix}.$$

The difference of the two vectors is

$$\vec{w} = \vec{u} - \vec{v} = \begin{bmatrix} u^x - v^x \\ u^y - v^y \\ u^z - v^z \end{bmatrix}.$$
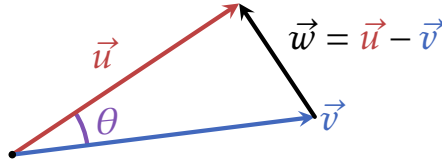
Using the law of cosines (**??**) we get

$$\left\| \vec{w} \right\| = \left\| \vec{u} \right\|^2 + \left\| \vec{v} \right\|^2 - 2 \left\| \vec{u} \right\| \left\| \vec{v} \right\| \cos(\theta).$$

Substituting in the definition of the dot product gives

$$\left\| \vec{w} \right\| = \left\| \vec{u} \right\|^2 + \left\| \vec{v} \right\|^2 - 2\vec{u} \cdot \vec{v},$$

which by rearrangement yields

$$\vec{u} \cdot \vec{v} = \frac{1}{2} \left(\left\| \vec{u} \right\|^2 + \left\| \vec{v} \right\|^2 - \left\| \vec{w} \right\| \right).$$

Now we simply substitute in the explicit forms of all the terms inside the paranthesis,

$$\vec{u} \cdot \vec{v} = \frac{1}{2} \left(\left(u^x \right)^2 + \left(u^y \right)^2 + \left(u^z \right)^2 + \left(v^x \right)^2 + \left(v^y \right)^2 + \left(v^z \right)^2 - \left(u^x - v^x \right)^2 - \left(u^y - v^y \right)^2 - \left(u^z - v^z \right)^2 \right),$$

expand and cancel terms. We finally get

$$\vec{u} \cdot \vec{v} = \frac{1}{\cancel{2}} \left(\cancel{2} u^x v^x + \cancel{2} u^y v^y + \cancel{2} u^z v^z \right) = u^x v^x + u^y v^y + u^z v^z. \tag{1.1.19}$$

---

[2]where's the fun otherwise?

Indeed quite a simple formula! If we use the general $\mathbb{R}^n$ instead of $\mathbb{R}^3$, we would get the general formula

$$\vec{u} \cdot \vec{v} = u^1 v^1 + u^2 v^2 + \cdots + u^n v^n = \sum_{k=1}^{n} u^k v^k. \tag{1.1.20}$$

> ❶ To write/do ❶
>
> examples?

A useful result of Equation 1.1.20 is that the dot product of a vector with itself gives the square of the vector's norm:

$$\vec{u} \cdot \vec{u} = \sum_{k=1}^{n} u^k u^k = \sum_{k=1}^{n} \left(u^k \right)^2 = \left\| \vec{u} \right\|^2. \tag{1.1.21}$$

> **Note 1.8 Row-column product**
>
> While Equation 1.1.20 is very neat, the way it is written here has some issues which are hard to explain at the moment - but are expanded on in later chapters. For now we just say that when we calculate the dot product of two vectors it is better to express the first one as a **row vector**. What precisely is a row vector and how it differs from a column vector is explained in the section **??**. For now just accept that a row vector is an object similar to a column vector, but we write its components using a subscript instead of a superscript. That is, the row vector $\underline{u}$ has components $u_1, u_2, \ldots, u_n$. We can then write
>
> $$\vec{u} \cdot \vec{v} \longrightarrow \underline{u} \cdot \vec{v} = \begin{bmatrix} u_1, & u_2, & \ldots, & u_n \end{bmatrix} \begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{bmatrix} = \sum_{k=1}^{n} u_k v^k = u_k v^k.$$
>
> (the last equality is using Einstein notation)
>
> So just keep this in mind. It might be a bit strange at the moment, but it will become more clear in later sections.

1.1.13 The cross product

Another commonly used product of two vectors is the so-called **cross product**. Unlike the dot product, it is only really valid in $\mathbb{R}^2$, $\mathbb{R}^3$ and $\mathbb{R}^7$, of which we will focus on $\mathbb{R}^3$ and touch a bit on its uses in $\mathbb{R}^2$. Also in contrast to the dot product, the cross product in $\mathbb{R}^3$ results in a vector rather than a scalar - therefore the product is sometimes known as the **vector product**. The cross product uses the notation $\vec{a} \times \vec{b}$, from which it derives its name.
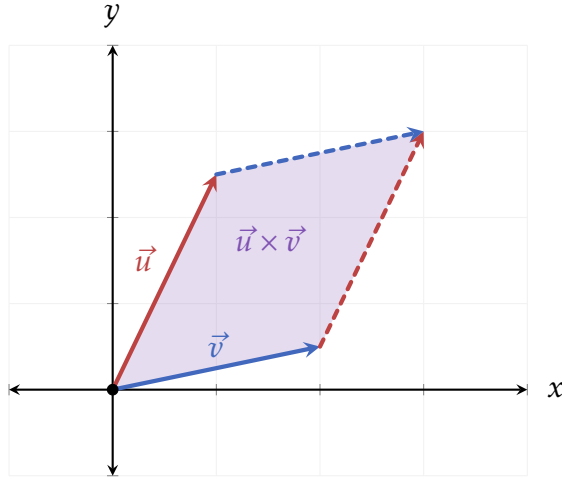
Figure 1.3 The cross product in $\mathbb{R}^2$ of two vectors $\vec{u} = \begin{bmatrix} a \\ b \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} c \\ d \end{bmatrix}$ as the signed area of the parallogram defined by the vectors.

We start with the definition of the cross product in $\mathbb{R}^2$: the cross product of two vectors $\vec{u} = \begin{bmatrix} a \\ b \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} c \\ d \end{bmatrix}$ is the (signed) area of the parallelogram defined by the two vectors (see Figure 1.3).

The value of the parallelogram defined by $\vec{u}$ and $\vec{v}$ is

$$\vec{u} \times \vec{v} = \left\| \vec{u} \right\| \left\| \vec{v} \right\| \sin(\theta), \tag{1.1.22}$$

where θ is the angle between the vectors. This is extremely similar to the scalar product, and we can use this fact to find how to calculate the cross product from vectors in column form: if we replace $\vec{u}$ by a vector orthogonal to it, denoted by $\vec{u}^\perp$, the cross product is then

$$\vec{u} \times \vec{v} = \left\| \vec{u}^\perp \right\| \left\| \vec{v} \right\| \sin\left(\theta + \frac{\pi}{2}\right), \tag{1.1.23}$$

since the angle between $\vec{u}^\perp$ and $\vec{v}$ is $\frac{\pi}{2}$ more than that between $\vec{u}$ and $\vec{v}$. Using the fact that $\sin\left(\theta + \frac{\pi}{2}\right) = \cos(\theta)$, we get the equality

$$\begin{aligned}
\vec{u} \times \vec{v} &= \left\| \vec{u}^\perp \right\| \left\| \vec{v} \right\| \sin\left(\theta + \frac{\pi}{2}\right) \\
&= \left\| \vec{u}^\perp \right\| \left\| \vec{v} \right\| \cos(\theta) \\
&= \vec{u}^\perp \cdot \vec{v}.
\end{aligned} \tag{1.1.24}$$

In $\mathbb{R}^2$, any vector $\vec{u} = \begin{bmatrix} a \\ b \end{bmatrix}$ has two vectors orthogonal to it: $\begin{bmatrix} -b \\ a \end{bmatrix}$ and $\begin{bmatrix} b \\ -a \end{bmatrix}$. Choosing the former gives

$$\vec{u} \times \vec{v} = \begin{bmatrix} -b \\ a \end{bmatrix} \cdot \begin{bmatrix} c \\ d \end{bmatrix} = -bc + ad, \tag{1.1.25}$$
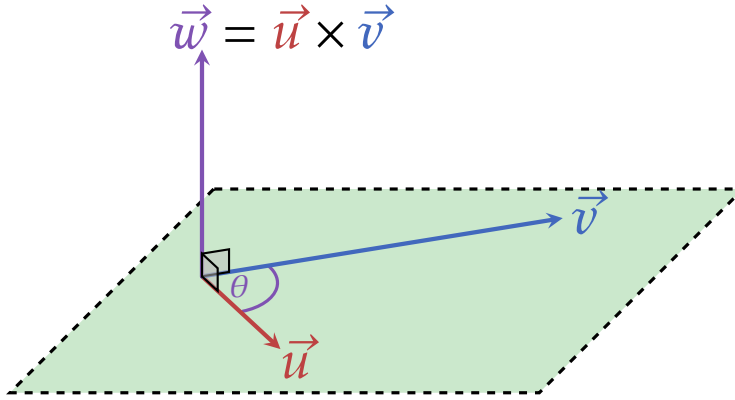
Figure 1.4 The cross product of the vectors $\vec{u}$ and $\vec{v}$ relative to the plane spanned by the two vectors.

while choosing the latter gives

$$\vec{u} \times \vec{v} = \begin{bmatrix} b \\ -a \end{bmatrix} \cdot \begin{bmatrix} c \\ d \end{bmatrix} = bc - ad. \tag{1.1.26}$$

These two forms are the opposite of each other - i.e. if one yields the value 4, the other yields the value -4. We will see which one is used in a moment.

On to $\mathbb{R}^3$: geometrically, the cross product of two vectors $\vec{u}, \vec{v} \in \mathbb{R}^3$ is defined as a **vector** $\vec{w} \in \mathbb{R}^3$ which is **orthogonal to both** $\vec{u}$ and $\vec{v}$, and with norm of the same magnitude as the product would have in $\mathbb{R}^2$, i.e.

$$\left\| \vec{w} \right\| = \left\| \vec{u} \right\| \left\| \vec{v} \right\| \sin(\theta). \tag{1.1.27}$$

The direction of $\vec{u} \times \vec{v}$ is determined by the **right-hand rule**: using a person's right hand, when $\vec{u}$ points in the direction of their index finger and $\vec{v}$ points in the direction of their middle finger, then vector $\vec{w} = \vec{u} \times \vec{v}$ points in the direction of their thumb:
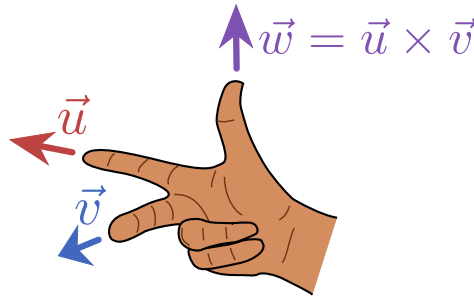
The cross product is **anti-commutative**, i.e. changing the order of the vectors results in inverting the product:

$$\vec{u} \times \vec{v} = -\left(\vec{v} \times \vec{u} \right).$$

When the vectors are given as column vectors, $\vec{u} = \begin{bmatrix} u^x \\ u^y \\ u^z \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} v^x \\ v^y \\ v^z \end{bmatrix}$, the resulting cross product is

$$\vec{u} \times \vec{v} = \begin{bmatrix} u^y v^z - u^z v^y \\ u^z v^x - u^x v^z \\ u^x v^y - u^y v^x \end{bmatrix}. \tag{1.1.28}$$

Note 1.9 The cross product of the standard basis vectors

The cross product of two of the standard basis vectors in $\mathbb{R}^3$ is the third basis vector. Its sign ($\pm$) is determined by a cyclic rule:

$$\text{sign}\left(\hat{e}_i \times \hat{e}_j\right) = \begin{cases} 1 & \text{if } (i,j) \in \left\{(1,2),\,(2,3),\,(3,1)\right\}, \\ -1 & \text{if } (i,j) \in \left\{(3,2),\,(2,1),\,(1,3)\right\}, \\ 0 & \text{otherwise.} \end{cases}$$

!

Challenge 1.1 Orthogonalily of the cross product

Using component calculation and utilizing the dot product, show that $\vec{a} \times \vec{v}$ is indeed orthogonal to both $\vec{a}$ and $\vec{b}$.

?

1.1.14 The Gram-Schmidt process

1.1.15 Normal vectors

1.1.16 Usage examples