

MATHEMATICS FOR SCIENCE STUDENTS

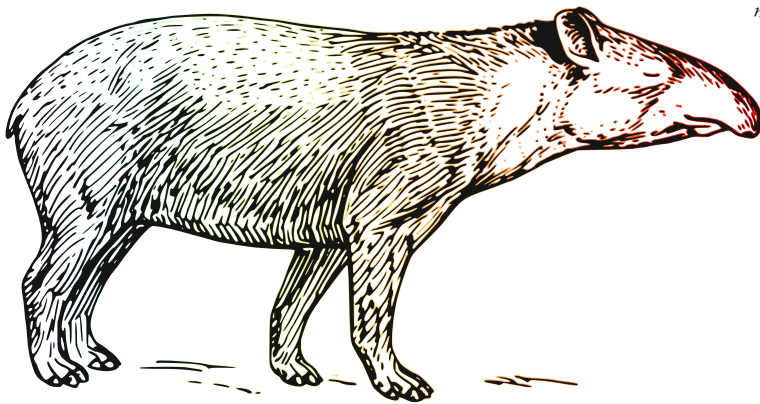
An open-source book

Written, illustrated and typeset (mostly) by

PELEG BAR SAPIR

with contributions from others

$$\begin{array}{l} a^b = e^{b \log(a)} \\ (a+b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \\ \binom{n}{k} = \frac{n!}{k!(n-k)!} \\ T(\alpha \vec{u} + \beta \vec{v}) = \alpha T(\vec{u}) + \beta T(\vec{v}) \\ \Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt \\ R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \\ A = Q \Lambda Q^{-1} \\ \frac{df}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \\ \langle \hat{e}_i, \hat{e}_j \rangle = \delta_{ij} \\ \vec{v} = \sum_{i=1}^n \alpha_i \hat{e}_i \\ e^{\pi i} + 1 = 0 \\ \int_a^b f(x) dx = F(b) - F(a) \\ \cos(x) = \sum_{n=0}^\infty \frac{(-1)^n}{(2n)!} x^{2n} \end{array}$$



PUBLISHED IN THE WILD

! To be written/to do: 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.

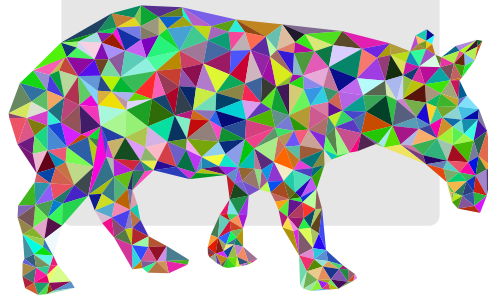
Note 0.1 In case you are already familiar with the topics

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 two 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

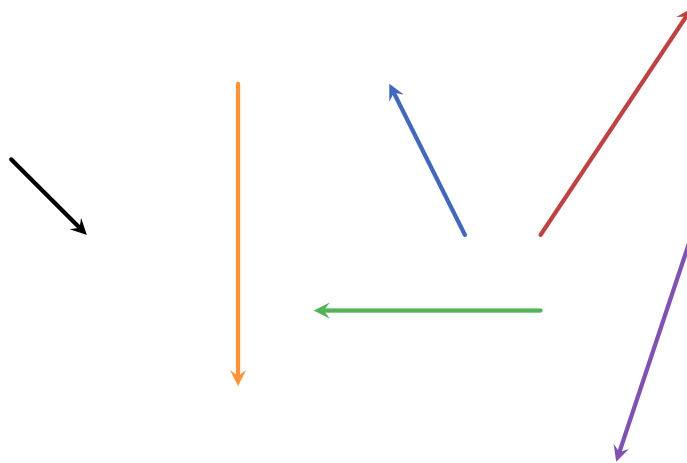
Vectors are the fundamental objects of linear algebra: the entire field revolves around manipulation of vectors. In this chapter we deal with the so-called **real vectors**, which can be defined in a geometric way:

Definition 1.1 Real vectors

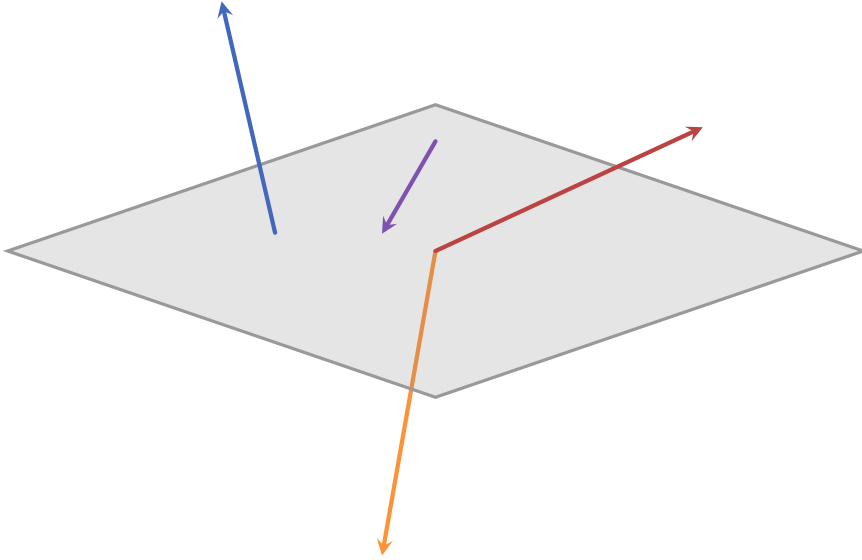
A *real vector* is an object with a **magnitude** (also called **norm**) and a **direction**.

π

In this chapter we refer to real vectors simply as *vectors*. Vectors can have 1, 2, 3, ... 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:



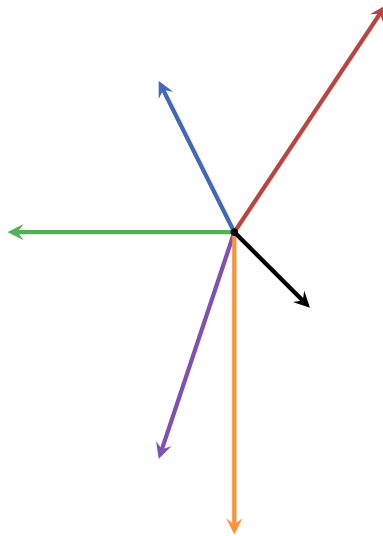
Unfortunately, it is difficult (if not impossible) to draw higher-dimensional vectors. For now, we will concentrate on 2-dimensional vectors and explore their properties. Later in the section we will apply what we learned from 2-dimensional vectors to 3-dimensional vectors and higher-dimensional vectors 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:** \mathbf{u} , \mathbf{v} , \mathbf{x} , \mathbf{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} : $\|\vec{v}\|$.

A vector \vec{v} with norm $\|\vec{v}\| = 1$ 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.

$$\hat{v} = \frac{1}{\|\vec{v}\|} \vec{v}. \quad (1.1.1)$$

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} .

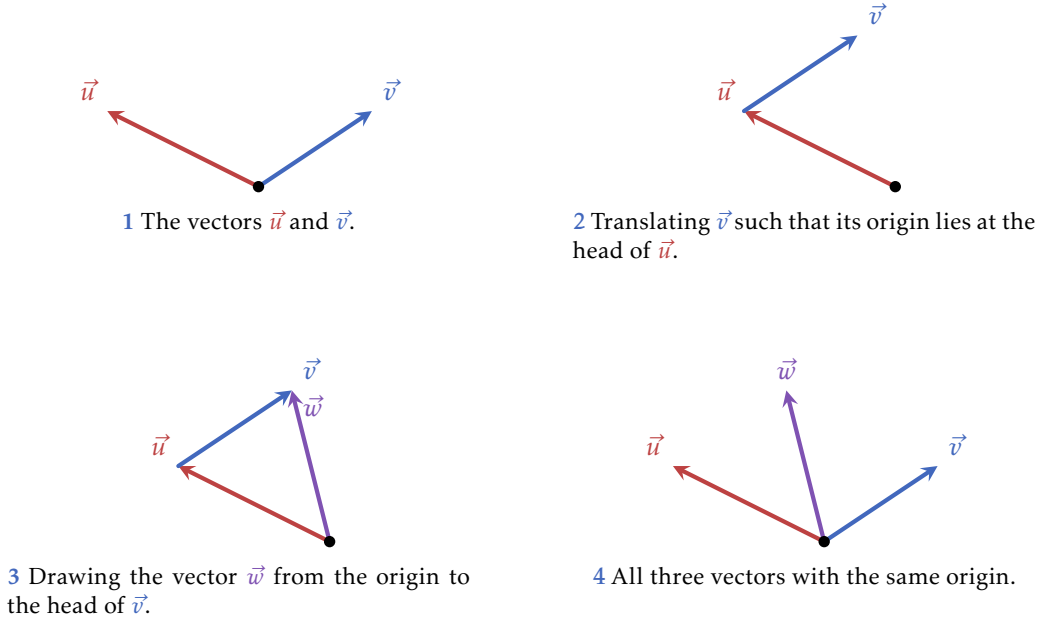


Figure 1.1 Vector addition.

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} ,

$$\vec{v} + \vec{0} = \vec{v}. \quad (1.1.2)$$

(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.

$$\vec{v} + (-\vec{v}) = \vec{0}. \quad (1.1.3)$$

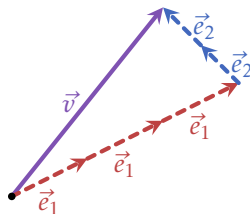
1.1.2 Column representation

In order to be able to use vectors for actual calculations we must somehow quantify their properties. When quantifying geometric shapes 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 a vector as a unit of measurement for each dimension (and call these vectors **basis vectors**). For example, we can choose the following two 2-dimensional vectors \vec{e}_1 and \vec{e}_2 to be used as basis vectors:



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

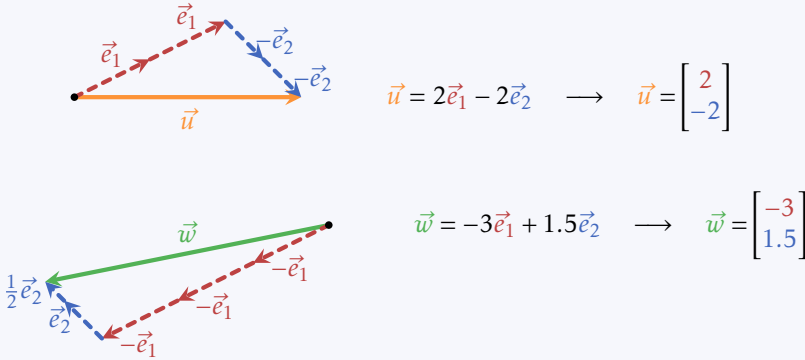


We see that using these basis vectors, $\vec{v} = 3\vec{e}_1 + 2\vec{e}_2$. This means that we need to add three times \vec{e}_1 and two times \vec{e}_2 to construct \vec{v} . We denote this fact by writing \vec{v} as a column of two numbers:

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

Example 1.1 Vectors from basis vectors

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



This notation is frequently referred to as a **column vector**, and the numbers are called **coordinates** or **components**. 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 $v^2 = 2$.

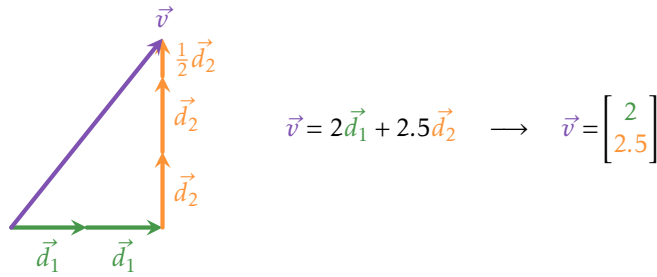
Note 1.2 Index notation

In this book we use upper-index notation to denote components of column vectors. Do not mistake these for powers! The reason for this choice (as opposed to low-index 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.

The set of basis vectors used to represent vectors as columns is sometimes called a **coordinate system**. We will see different common coordinate systems soon. Vectors have different components in different coordinate systems. For example, we can use the following two vectors \vec{d}_1 and \vec{d}_2 as basis vectors:



In this new coordinate system, the vector \vec{v} from earlier has the following column representation:



(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 coordinate system we use, vectors remain the same - it's their *representation* which changes.

In fact, not only does the choice of coordinate system affect the column representation of all vectors¹, two different vectors **can have the same column representation using two different coordinate systems..** 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:

$$\begin{aligned}\vec{u} &= 1 \cdot \vec{u} + 0 \cdot \vec{v}, \\ \vec{v} &= 0 \cdot \vec{u} + 1 \cdot \vec{v}.\end{aligned}$$

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.

! To be written/to do: more examples? !

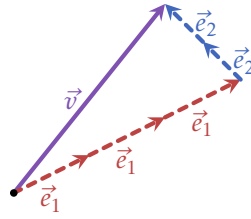
1.1.3 Vector operations, norm and the zero vector in column form

Previously we saw how to scale and add vectors (the latter using the parallelogram 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:

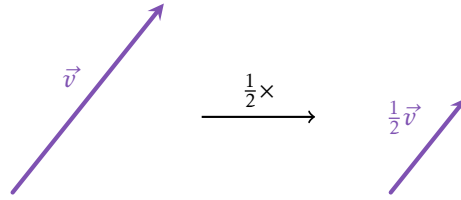


¹except the zero vector, which is always $\vec{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

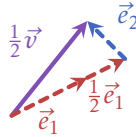
We start with the vector \vec{v} from before, i.e. $3 \cdot \vec{e}_1 + 2 \cdot \vec{e}_2$ and $\vec{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$:



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 $\frac{1}{2}\vec{v}$:

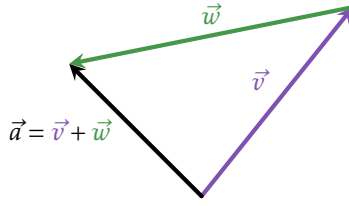


We get that $\frac{1}{2}\vec{v} = 1.5\vec{e}_1 + 1\vec{e}_2$, i.e. $\frac{1}{2}\vec{v} = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \cdot 3 \\ \frac{1}{2} \cdot 2 \end{bmatrix}$ - i.e. scaling \vec{v} by $\frac{1}{2}$ simply multiplied both its components by $\frac{1}{2}$.

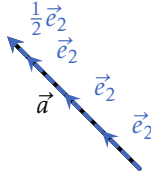
This is in fact true for any vector and any scalar, using any coordinate system: scaling a vector by a scalar α results in multiplying each of its components by α :

$$\vec{u} = \begin{bmatrix} u^1 \\ u^2 \end{bmatrix} \implies \alpha \vec{u} = \begin{bmatrix} \alpha \cdot u^1 \\ \alpha \cdot u^2 \end{bmatrix}. \quad (1.1.4)$$

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 always true:

$$\vec{u} = \begin{bmatrix} u^1 \\ u^2 \end{bmatrix}, \vec{v} = \begin{bmatrix} v^1 \\ v^2 \end{bmatrix} \implies \vec{u} + \vec{v} = \begin{bmatrix} u^1 + v^1 \\ u^2 + v^2 \end{bmatrix}. \quad (1.1.5)$$

Note 1.3 Scaling and addition using different coordinate systems

This is only true if the column representation of the two vectors is in the same coordinate system (i.e. using the same basis vectors). Adding two column vectors in different coordinate systems (i.e. using different basis vectors) requires changing one of the column representation to the basis of the other one.



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

$$-\vec{v} = \begin{bmatrix} -v^1 \\ -v^2 \end{bmatrix}. \quad (1.1.6)$$

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

$$\vec{0} = \begin{bmatrix} v^1 \\ v^2 \end{bmatrix} + \left(-1 \begin{bmatrix} v^1 \\ v^2 \end{bmatrix} \right) = \begin{bmatrix} v^1 \\ v^2 \end{bmatrix} + \begin{bmatrix} -v^1 \\ -v^2 \end{bmatrix} = \begin{bmatrix} v^1 - v^1 \\ v^2 - v^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (1.1.7)$$

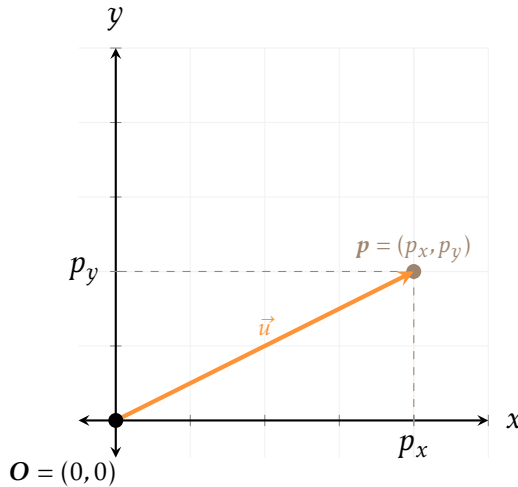
Summary 1.1 Column representation of vectors

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 The space \mathbb{R}^2 and the standard basis set**

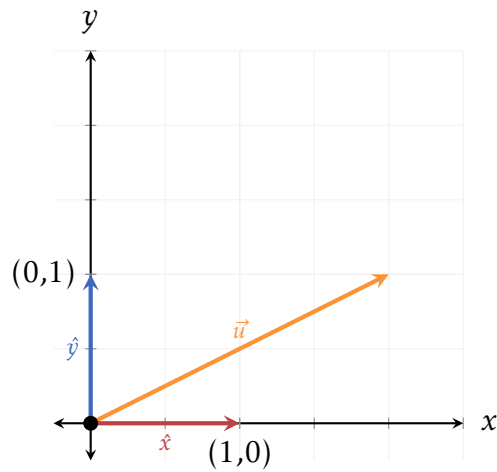
We can place vectors on a two-dimensional Cartesian coordinate system, such that their origin coincide with the axis-origin (the point $O = (0, 0)$). We then mark the point where its head lies as $p = (p_x, p_y)$:



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 represented $\vec{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.