MATHEMATICS FOR SCIENCE STUDENTS

An open-source book

Written, illustrated and typeset (mostly) by

PELEG BAR SAPIR

with contributions from others

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

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

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

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

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

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

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

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

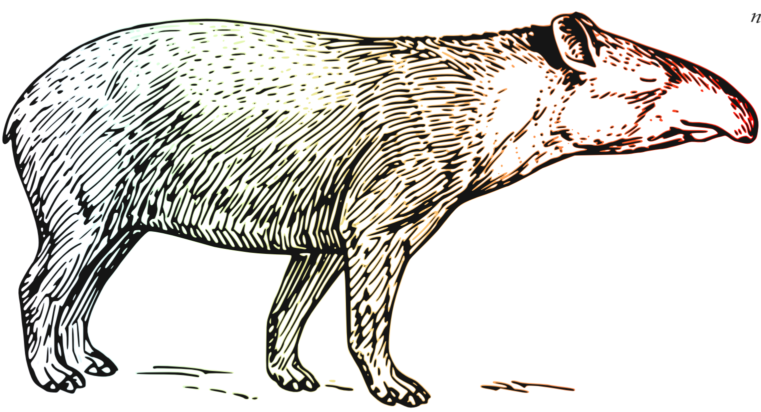
$\langle \hat{e}_i, \hat{e}_j \rangle = \delta_{ij}$

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

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

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

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



PUBLISHED IN THE WILD

! **To be written**: Rights, lefts, etc. will be written here in the future !

HERE BE TABLE

CHAPTER

0

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.

0.1 FUNCTIONS AS VECTORS

Note 0.1 Calculus ahead!

This section uses ideas discussed in the chapter about 1-dimensional real calculus (**??**). While strict knowledge of calculus is not necessary, some concepts would be rather difficult to understand without it.

!

Up until now we used a rather informal definition for vectors (**??**). While the formal definition will be discussed in the next chapter, it is worth while to review the basic properties of vectors we saw so far. We start with the scaling of vectors:

- **It is 'close'**: the result of scaling a vector is itself a vector.
- **The scalar 1 is neutral to scaling**: i.e. $1\cdot\vec{v}=\vec{v}$.
- **It is associative**: for any $\alpha,\beta\in\mathbb{R}$ and $\vec{v}\in\mathbb{R}^{n}$,

$$\alpha\cdot\left(\beta\vec{v}\right)=\left(\alpha\cdot\beta\right)\vec{v}.$$

And for vector addition:

- **It is close**: the sum of any two vectors is also a vector.
- **It is commutative**: the order of addition does not change its result.
- **It is associative**: while not discussed directly, it is obvious that when adding together any three vectors $\vec{u},\vec{v},\vec{w}\in\mathbb{R}^{n}$, we can change the order of addition and that too will not change its result: calculating $\vec{u}+\vec{v}$ first and then adding $\vec{w}$ to the result is the same as calculating $\vec{v}+\vec{w}$ first and then adding $\vec{u}$ to the result, i.e.

$$\left(\vec{u}+\vec{v}\right)+\vec{w}=\vec{u}+\left(\vec{v}+\vec{w}\right).$$

- $\vec{\mathbf{0}}$ **is neutral to addition**: i.e. the addition of $\vec{0}$ to any vector $\vec{v}$ results in $\vec{v}$.

- **Any vector has an additive inverse**: given a vector $\vec{u} \in \mathbb{R}^n$, there is always a vector $\vec{v} \in \mathbb{R}^n$ such that

$$\vec{u} + \vec{v} = \vec{0},$$

namely the vector $\vec{u} = -\vec{v}$.

The two operations also have two important properties together:

- **Vector addition is distributive**: for any $\alpha \in \mathbb{R}$ and $\vec{u}, \vec{v} \in \mathbb{R}^n$,

$$\alpha\left(\vec{u} + \vec{v}\right) = \alpha\vec{u} + \alpha\vec{v}.$$

- **Scalar multiplication is distributive**: for any $\alpha, \beta \in \mathbb{R}$ and $\vec{v} \in \mathbb{R}^n$,

$$\left(\alpha + \beta\right)\vec{v} = \alpha\vec{v} + \beta\vec{v}.$$

All these properties seem rather obvious, but not all mathematical structures actually have them: for example, in a later chapter in the book we learn about **groups**, which do not have a scaling operator and don't necessarily have some of these properties, e.g. not all of them are commutative under their own operation.

Real functions, on the other hand, do have all these properties. We can therefore apply to real functions all the stuff we learned about vectors in this chapter. While we're not promised that everything will *look* the same, their general behavior is identical to that of vectors: the two obvious examples are scaling and addition:

- **Scaling**: given a real function f we can always scale it by multiplying its output by any real number α: if $f(x) = y$, then

$$\alpha f(x) = \alpha y.$$

For example, we can scale the polynomial $P(x) = x^3 - 2x + 5$ by a factor of 7, yielding

$$7P(x) = 7 \cdot x^3 - 7 \cdot 2x + 7 \cdot 5 = 7x^3 - 14x + 35.$$

And indeed we see that the result is itself a real function.

- **Addition**: any two real functions f, g can be added together. For example, given the functions $f(x) = \mathrm{e}^x$ and $g(x) = 5\sin(x)$, their sum is then

$$\left[f + g\right](x) = \mathrm{e}^x + 5\sin(x),$$

which is indeed a real function by itself.

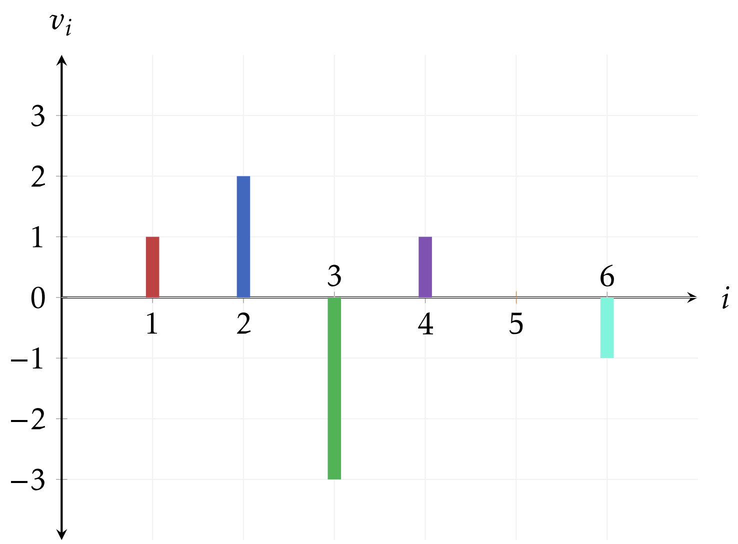
You should go over all the above properties of vectors and verify for yourself that real functions do indeed posses them.

Before we move on, we should limit our further discussion to a specific set of real functions in order to avoid some annoying details which might rise later, and are not critical for the understanding of the idea of functions as vectors. Thus, from now on in the section we will only discuss real functions which are continuous in all of $\mathbb{R}$, and are "smooth" enough to not cause too many problems. Recall from **??** that by a "smooth" function we mean a function infinitely continuous derivatives (i.e. its derivative of any order is continuous).

The first important thing to do when using vectors is to choose a basis set to represent them. For any $\mathbb{R}^n$ it is rather easy to understand how this representation looks like: we simply write the vector with n components. How can we do that with functions? Well, consider the following vector in $\mathbb{R}^6$:

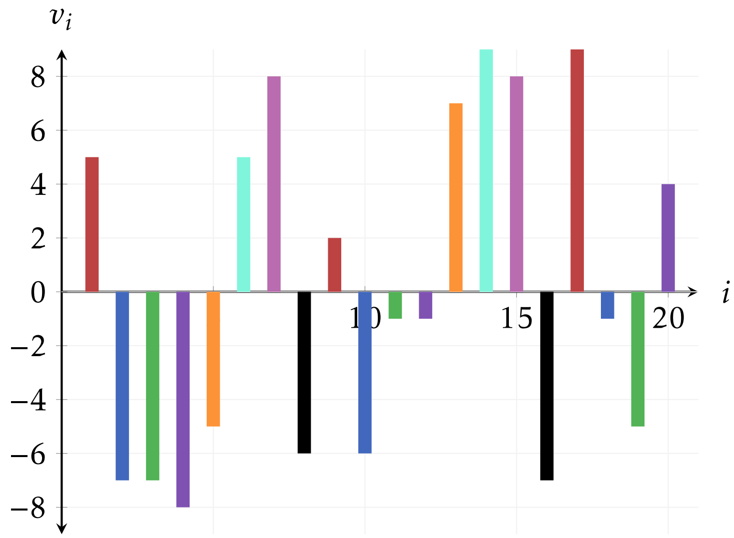
$$\vec{v} = \begin{bmatrix} 1 \\ 2 \\ -3 \\ 1 \\ 0 \\ -1 \end{bmatrix}. \tag{0.1.1}$$

While drawing 6-dimensional spaces is rather difficult, we can draw as a bar chart each of its components v_i as a function of the component index i:



We can do the same for the $\mathbb{R}^{20}$ vector

$$\vec{u} = \left[5,\ -7,\ -7,\ -8,\ -5,\ 5,\ 8,\ -6,\ 2,\ -6,\ -1,\ -1,\ 7,\ 9,\ 8,\ -7,\ 9,\ -1,\ -5,\ 4\right]^\top : \tag{0.1.2}$$



We can continue this method for any natural number, or even an integer. In fact, we can expand it to any real number: given a number $x \in \mathbb{R}$, we assign it some number $y \in \mathbb{R}$ and say that y is the component corresponding to the index x... and we just defined a real function!