

# Notes on Linear Algebra I. # 1

Oussama Moutaouikil

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>On Vector Spaces</b>	<b>5</b>
2.1	Vectors . . . . .	5
2.2	Vector Spaces . . . . .	7
2.3	A Good Stack of Examples . . . . .	7
2.4	Subspaces . . . . .	7
2.5	Sums and Direct Sums . . . . .	7
2.6	Linear Combinations . . . . .	7
2.7	Span . . . . .	7

## 1 Introduction

*Linear algebra* is the study of *linear transformations* on *finite-dimensional vector spaces*. All these words may seem unfamiliar to you for now but they will become clear once we deal with a great stack of examples.

A transformation is any operation that takes an input and transforms it into an output. There are many types of transformations, but only some of them are *linear*. This last class of transformations is the one that satisfy the following properties :

1. If you amplify the input then there will be a corresponding amplification to the output. (e.g. tripling the input will triple the output) This property is known as *homogeneity*.
2. Adding inputs together leads to adding of their respective outputs. This is known as *additivity*.

An example of such transformation is the map (or function)  $x \mapsto 10x$ . It is linear since if you amplify the input by some number  $a$ , then the output will also be amplified by  $a$ , i.e.  $ax \mapsto a(10x)$ . And let's say that for some  $x_1$  the output is  $10x_1$ , and for some  $x_2$  the output is  $10x_2$ , then adding those two inputs  $x_1 + x_2$  leads to adding of their respective outputs  $10x_1 + 10x_2$  :

$$x_1 + x_2 \mapsto 10(x_1 + x_2) = 10x_1 + 10x_2.$$

Input $x$	Transformation	Output $10x$
2	$\rightarrow$	20
4	$\rightarrow$	40
$2 + 4 = 6$	$\rightarrow$	$60 = 20 + 40$
$5 \cdot 2 = 10$	$\rightarrow$	$100 = 5 \cdot 20$

In general, any map of the form  $x \mapsto cx$ , where  $c$  is some constant, is a linear transformation as you can check.

A non-example of a linear transformation is the map  $x \mapsto x^2$ , doubling the input quadruples the output since  $2x \mapsto (2x)^2 = 4x^2$ . Also if one adds two inputs together, their outputs do not add, example: an input of 2 has a 4 output, and an input of 3 has a 9 output. If it was linear one would expect that the input  $2 + 3 = 5$  has a  $4 + 9 = 13$  output, yet if we input 5 we get  $5^2 = 25$ !

Input $x$	Transformation	Output $x^2$
2	$\rightarrow$	4
3	$\rightarrow$	9
$2 + 3 = 5$	$\rightarrow$	$25 \neq 13 = 4 + 9$
$2 \cdot 3 = 6$	$\rightarrow$	$36 \neq 18 = 2 \cdot 9$

If we now graph each map in an  $xy$ -coordinate system we can see that our first map represents a line, whereas the second is highly non-linear :

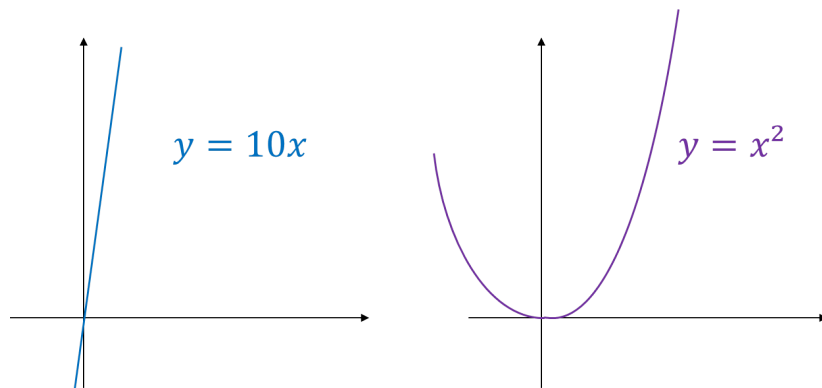


Figure 1: Comparison between the graph of  $x \mapsto 10x$  and of  $x \mapsto x^2$ .

Now we may ask: “If the graph of a certain transformation is a line, does it mean that it is linear?” Well we know that the equation of a line is  $y = ax + b$ , i.e. the map  $x \mapsto ax + b$ , where  $a, b$  are constants in  $\mathbf{R}$ . For it to be a linear transformation, a doubling in the input should produce a doubling of the output, so we double the input  $x$  to get  $2x$ . If we apply the transformation we get  $2ax + b$ . However, if we double the output  $y$  we get  $2(ax + b) = 2ax + 2b$ , those are equal if and only if :

$$2ax + b = 2ax + 2b \iff b = 2b \iff b = 0.$$

What this tells us is that the line must have as an equation  $y = ax$ , hence it must pass through the origin since the coordinates of the point  $(0, 0)$  satisfies that equation. So if the graph of a certain transformation is a line doesn't necessarily mean that it is linear, it needs also to pass through the origin!

A lot of mathematical structures exhibit linear properties. Let's recall for instance what we learnt in Calculus. Say  $f, g, h$  are some differentiable functions on a nonempty interval  $I \subseteq \mathbf{R}$ , and that  $f(x) = g(x) + h(x)$  for all  $x \in I$  (in other words the function  $f$  can be expressed as a sum of two other functions), then :

$$\frac{d}{dx}f(x) = f'(x) = \frac{d}{dx}[g(x) + h(x)] = g'(x) + h'(x).$$

We wish to write this fact using a more compact notation. To do so we consider the *differential operator*  $\tilde{D}_x$ , which is a transformation from a function space  $\mathcal{F}_1$  to another function space  $\mathcal{F}_2$ . It takes as an input a function and outputs its derivative. (A function space is just a set of functions.) So instead of writing  $f'(x)$  or  $\frac{d}{dx}$  we can simply write  $\tilde{D}_x[f]$ . Hence we can write the previous identity as :

$$\tilde{D}_x[f] = \tilde{D}_x[g + h] = \tilde{D}_x[g] + \tilde{D}_x[h].$$

Now let  $c$  be a constant in  $\mathbf{R}$ . We define the function  $cf$  as  $(cf)(x) = cf(x)$ . Then its derivative is :

$$\frac{d}{dx}(cf)(x) = \frac{d}{dx}cf(x) = c\frac{d}{dx}f(x) = cf'(x).$$

Which we can write using the differential operator :

$$\tilde{D}_x[cf] = c\tilde{D}_x[f].$$

In other examples the transformation acted on a number and outputed a number. Now the differential operator acts on functions and outputs function. Yet it also shows signs of linearity, so for example a doubling in the input causes a doubling in the output. And adding inputs together leads to adding of their respective outputs. Note also that the “zero function”  $0(x) : x \mapsto 0$  can be part of the function space where the operator can act, since it is differentiable.

Indeed, even the point of differential calculus is to approximate non-linear functions using linear functions (e.g. approximating a curve using a tangent line), so it is of no surprise to see these rules built into the core of Calculus. It will also be unsurprising to find that linear algebra can help us in solving differential equations.

Now what if we think of a transformation that acts on *vectors*? (Informally, a vector is a quantity that can be represented using a magnitude and a direction, although this is sort of unrigorous and is mainly true for some physical vectors – and is false for some other kinds. They are in most cases denoted using a bold font like  $\mathbf{N}$ , or with an arrow in the top  $\vec{N}$  to distinguish them from scalars) Hooke's law is a principle that states that the force  $\vec{F}$  needed to extend or compress a spring by some distance  $\vec{X}$  is proportional to that distance. In other words, there is a constant  $k$  such that  $\vec{F} = k\vec{X}$ , which can be rewritten as  $\vec{X} = \vec{F}/k$ . Note that  $\vec{F}$  and  $\vec{X}$  are vector quantities. In our example we let  $k = 5 \text{ N} \cdot \text{m}^{-1}$ . So if you extend a spring (see Figure 2 to visualize more the direction of the force) using a downward force of 5 N, then the displacement  $\vec{F}$  would be 25 m downward. And if you compress the spring by an upward force of 2 N, then the resultant displacement would be 10 m. Adding the inputs 5 N downward plus 2 N upward (which is 3 N downward) produces as an output 25 m downward plus 10 m upward (which is 15 m downward). As you can check, it satisfies linearity.

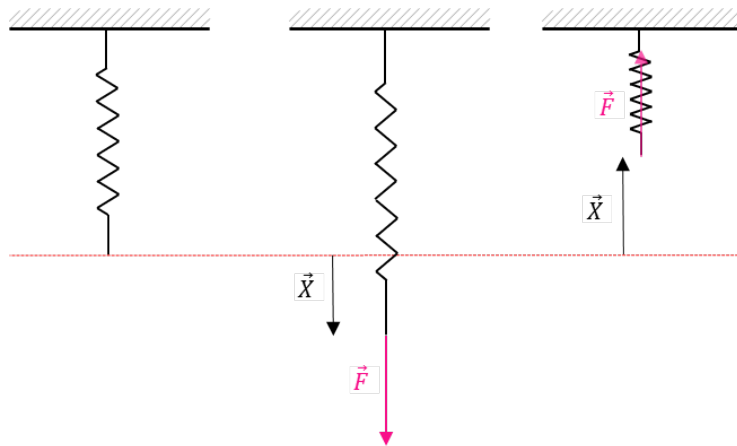


Figure 2: Hooke's law  $\vec{F} = k\vec{X}$  in action.

If you apply a very large force which exceeds the limit (since some permanent deformation or change of state once the spring is stretched or compressed to a certain point) then Hooke's no longer holds. However Hooke's law does hold for small forces and deformations, so it is a good approximation for most solid bodies.

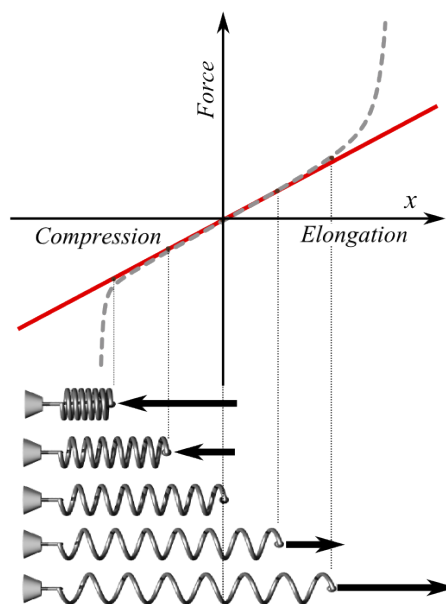


Figure 3: Applied force  $\vec{F}$  vs. elongation  $\vec{X}$  according to Hooke's law (red line) and what the actual plot might look like (dashed line).

This is a good example since it shows that in real life things can be non-linear, however we can make good linear approximations to them which is useful.

We will also deal with matrices, those are mathematical objects like :

$$\begin{pmatrix} 1 & 0 & -1 \\ 93 & \pi^e & 20 \\ 7 & 9 & 9 \end{pmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}$$

which are also related to linear transformations. This relation as well as a rigorous definition of a matrix will be given later.

To summarize, linear algebra deals with the algebra of linear transformations on finite-dimensional vector spaces. The algebra that you may have learned deals with equating scalars by solving equations, simplifying expressions ... and performing operations on them, the same thing can be said about linear algebra where we'll extend the range of operations to a new family made of : direct sums, span, dimension, transpose, determinant, trace, eigenvalue, eigenvector, and characteristic polynomial. Those will be used to manipulate not only scalars, but also vectors, vector spaces, subspaces, inner spaces, ...

## 2 On Vector Spaces

### 2.1 Vectors

We've discussed in the introduction scalars which are quantities that can be described using only a single number. Examples of scalars are temperature, mass, volume, area, charge, density, interest, populations number, work, energy, ... We'll deal with many sets of scalars amongst them are  $\mathbf{R}$  and  $\mathbf{C}$ . Now some of you may not be familiar with the set  $\mathbf{C}$  (which stands for complex numbers), so we'll introduce it now.

After the formula for the quadratic equation  $ax^2 + bx + c = 0$  was discovered, which is :

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \quad a \neq 0$$

mathematicians began to search for a similar formula for the cubic equation  $ax^3 + bx^2 + cx + d = 0$ . We can show, using a tricky substitution, that this equation is equivalent to :

$$x^3 = 3px + 2q.$$

Cardano showed that this equation could be solved by the formula :

$$x = \sqrt[3]{q + \sqrt{q^2 - p^3}} + \sqrt[3]{q - \sqrt{q^2 - p^3}}.$$

Let's apply this formula to the cubic equation  $x^3 = 15x + 4$ . But first let's remind ourselves that the solution to that polynomial equation represents the intersection of the curve  $y = x^3$  and the line  $y = 15x + 4$ . We give a graph of it in Figure 4.

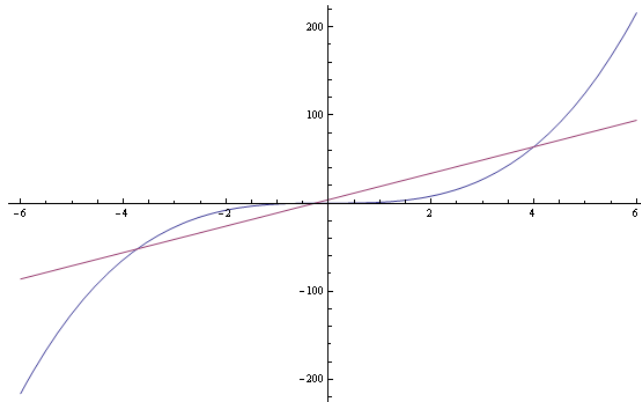


Figure 4: Plot of  $y = 15x + 4$  and of  $y = x^3$ .

We see that the line and the curve do meet on 3 occasions, so there are 3 solutions to our equation. To determine it we'll use Cardano's formula, we get :

$$x = \sqrt[3]{2 + \sqrt{-11}} + \sqrt[3]{2 - \sqrt{-11}}.$$

But  $\sqrt{-11}$  doesn't make sense. Yet from the graph we know that there are 3 solutions! This "paradox" was the source of Bombelli's "wild thought". He defined a new class of numbers where the square root of a negative number would make sense. Those are the numbers which form an ordered pair  $(a, b)$ , where  $a, b \in \mathbf{R}$ , but you can write this as  $a + bi$ . Those ordered pairs have a particular arithmetic, addition is defined as :

$$(a + bi) + (c + di) = (a + c) + (b + d)i,$$

and multiplication is defined by

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i;$$

where  $a, b, c, d \in \mathbf{R}$  Using this definition you can check that  $(0 + 1i)(0 + 1i) = i^2 = -1$ , from which we can say that  $i = \sqrt{-1}$ . You may also note that  $\mathbf{R}$  is in some sort a subset of  $\mathbf{C}$  if you consider  $a + 0i$ , so we have the hierarchy :

$$\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R} \subset \mathbf{C}.$$

Anyway using these new numbers we can derive all the real solutions to the equation we encountered, and this is one of their useful applications. They also have applications in multiple areas in physics. These numbers are called complex numbers, they share the same properties as those of real numbers :

#### Commutativity

$$z + w = w + z \text{ and } zw = wz \text{ for all } z, w \in \mathbf{C}.$$

#### Associativity

$$(z + v) + w = z + (v + w) \text{ and } (zv)w = z(vw) \text{ for all } z, v, w \in \mathbf{C}.$$

#### Existence of Identities

$$z + 0 = z \text{ and } z \cdot 1 = z \text{ for all } z \in \mathbf{C}.$$

### Additive Inverses

For every  $z \in \mathbf{C}$  there exists a unique  $w \in \mathbf{C}$  such that  $z + w = 0$ . The additive inverse of  $z$  is denoted as  $-z$ .

### Multiplicative Inverses

If  $z \in \mathbf{C}$  such that  $z \neq 0$ , then there exists a unique  $w \in \mathbf{C}$  such that  $zw = 1$ . The multiplicative inverse of  $z$  is denoted as  $1/z$ .

### Distributive Property

$z(v + w) = zv + zw$  for all  $z, v, w \in \mathbf{C}$ .

Both  $\mathbf{R}$  and  $\mathbf{Q}$  form what is called a *field*, a set along with two binary operations which satisfy the same properties that hold for  $\mathbf{R}$  and  $\mathbf{C}$ , namely commutativity, associativity, existence of identities, ...

We now discuss Cartesian products of sets. As an example  $\mathbf{R} \times \mathbf{R} = \mathbf{R}^2$  is defined to be the set of all ordered pairs  $(x, y)$  such that  $x, y \in \mathbf{R}$ . (Don't confuse it with complex numbers, remember  $\mathbf{C}$  has a different operation of multiplication) You can picture  $\mathbf{R}^2$  as a plane.

$$\mathbf{R}^2 = \{(x, y) : x, y \in \mathbf{R}\}.$$

What about  $\mathbf{R} \times \mathbf{R} \times \mathbf{R} = \mathbf{R}^3$ ? Well it is made of all ordered triplets of reals :

$$\mathbf{R}^3 = \{(x, y, z) : x, y, z \in \mathbf{R}\}.$$

So we can picture  $\mathbf{R}^3$  as ordinary 3D space. In general you see that

$$\underbrace{\mathbf{R} \times \dots \times \mathbf{R}}_{n \text{ times}} = \mathbf{R}^n, \quad n \in \mathbf{N}_{>0},$$

is the set of all ordered  $n$ -tuples  $(x_1, \dots, x_n)$  of real numbers. Now what if instead of real numbers we pick complex numbers? Well let's start from the basic  $\mathbf{C} \times \mathbf{C} = \mathbf{C}^2$ , it is made of all ordered pairs  $(z_1, z_2)$  of complex numbers. What about  $\mathbf{C}^3$ ? As you guessed it, it's the set of all ordered triples  $(z_1, z_2, z_3)$

## 2.2 Vector Spaces

In the previous section we made some discussions about vectors.

### 2.3 A Good Stack of Examples

#### 2.4 Subspaces

#### 2.5 Sums and Direct Sums

#### 2.6 Linear Combinations

#### 2.7 Span