1 Why mathematics?

The book of Nature is written in the language of Mathematics

Galileo Galilei

The unreasonable effectiveness of mathematics in the natural sciences

Eugene Wigner

These two statements reflect the different mindset of 16th and 20th century, and the different state of mathematics as a tool to understand the natural world in same eras. In Galieo's time it was common to think that the world had been designed by an intellect similar, although superior, to the human one. Galileo wanted to suggest that such a designer had used the language of Mathematics, and proposed using such a language to study the "Natural Philosophy". In such a way Physics was born.

More than 300 years later, mathematics had reached a great ability to describe the natural world. Physicists in Wigner's generation were discovering that ideas developed by pure mathematicians in abstract problems (e.g., Complex Numbers) found direct applications in new realms of Physics. It almost looked like the mathematicians' speculations were telling to the world how to behave. Particles seem to exist if and only if some mathematical structure allowed them to exist. This looked uncanny to their secular mind.

But if you are not a philosopher of science, and you just want to "act effectively" on the world as an engineer, you may be content to know that the mathematical language is indeed extremely useful and effective to do it.

Mathematical ideas resulted to be extremely effective to model physical phenomena, in particular regarding general laws and the behaviour of elementary entities. When systems become more complex, such as in real world applications, it is often impossible to find an exact solution as those studied in elementary calculus, but mathematics still allows us to formulate the problem in a rigorous manner and provides methods (algorithms) to find approximate but reliable solutions.

Indeed nowadays large part of human knowledge is expressed in mathematical form, although its mathematical nature is often hidden to the end user (e.g. modern operating systems).

Your relationship with mathematics "as an end user" depends on the nature of your work. A professional mathematician is supposed to develop novel mathematical ideas and results. An applied mathematician, a physicist, a researcher in Natural Sciences, a research engineer, a theoretical economist, etc., are supposed to be familiar with fairly advanced mathematical concepts and find new way to apply them to real world theoretical or practical problems.

For most engineering works it is enough to be familiar with current solution methods for practical problems. Such solutions are usually nowadays implemented as computer software, and a popular software for solving mathematical (numerical, i.e. by using algorithms for approximate solutions) problems is Matlab, that we are going to use in this course.

It is nevertheless important to understand the mathematical nature of the problem one has to solve (otherwise not only we won't understand the solution, but we will not even know that a solution exists, or how to formulate the problem). The time in which "Since now we have computers we do not need to study mathematics" is still far, and when it will come the human brain will be probably surpassed for the solution of any kind of problem, not only mathematical ones.

1.1 An example of modelling and numerical solution

Finding the equilibrium between two systems, one decreasing in a way proportional to its size, and the other oscillating, or finding the solutions of

$$e^{-t} = \sin t. \tag{1}$$

2 Vectors and matrices

Matlab stands for "matrix laboratory", and as the name says matrices are extremely important in the software. This is not surprising, since matrices are invaluable tools for mathematically operating in and on the natural world. Our first task will thus be not only to get familiar with the basic notions regarding matrices, but also with the *reason why* they are important.

2.1 Vectors: why

If you are located in the middle of a room, you may reach any point of the room by walking on a straight line (ignoring the possible presence of walls). Once the line is decided, though, the movement along that particular line does not allow you to reach any point in the room. You need to choose at least another direction not parallel to the first. When such two directions are decided, you may connect any two points on the floor by moving along them in an alternate way. For this reason the floor is called a $2D$ surface. But if you want to move to another floor, you will need again a new direction, not lying on the floor's plane. You may choose, for example (but not necessarily) a direction perpendicular to the floor, and you will be able, by combining movements in the 3 directions (and again ignoring obstacles), each point in space. No 4th direction is needed, and that is why we say the space we live in is $3D$.

The formalisation and generalisation of these ideas are the object of study of the theory of vectors and matrices, and are obviously important in describing the real world.

If we choose 3 perpendicular directions ($\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, $\hat{\mathbf{k}}$), we may identify each point in space with respect to a chosen *origin* using 3 numbers (coordinates), denoting the distances (in a proper unit) to be walked along the 3 chosen directions in order to reach the point (where walking in the opposite direction with respect to the chosen one is identified by a negative number), such as

$$\mathbf{x} \equiv \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \text{ e.g. } \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}. \tag{2}$$

Since any two points in space may be connected using a straight line, the vector $\mathbf{x}$ is also a direction in space. Its length is given by Pythagoras' theorem

$$l(\mathbf{x}) = \sqrt{x_1^2 + x_2^2 + x_3^2}. \tag{3}$$

We may thus call the movement along the direction of $\mathbf{x}$ for a length $c\,l(\mathbf{x})$ a new vector

$$z \equiv c\,\mathbf{x} = \begin{pmatrix} c\,x_1 \\ c\,x_2 \\ c\,x_3 \end{pmatrix}. \tag{4}$$

In the same way, by moving first along the direction of $\mathbf{x}$ for $l(\mathbf{x})$ and then along the direction of $\mathbf{y}$ for $l(\mathbf{y})$ we find ourselves in

$$\mathbf{z} \equiv \mathbf{x} + \mathbf{y} = \begin{pmatrix} z_1 = x_1 + y_1 \\ z_2 = x_2 + y_2 \\ z_3 = x_3 + y_3 \end{pmatrix} \tag{5}$$

with length

$$l(\mathbf{z}) = \sqrt{z_1^2 + z_2^2 + z_3^2} \leq l(\mathbf{x}) + l(\mathbf{y}). \tag{6}$$

The motion of physical bodies can be described studying how their position changes depending on a single parameter (real number) t, time, $\mathbf{x}(t)$, i.e. by using 3 real functions, one for each coordinate. Their time derivatives (again 3 functions for 3 coordinates) gives the body's velocity, and by derivation again we obtain the acceleration vector, which is connected through Newton's second law

$$m\mathbf{a} = \mathbf{F} \tag{7}$$

to the forces that determine the body's dynamics. Experiments tell us that forces add as vectors, and thus the $3D$ vector structure is fundamental to describe the motion of bodies.

2.2 Vector: generalisation

Should we be content with 3D vectors, since space is $3D$? Modern physics actually describes the world using $4D$ vectors (since time and space are not independent but somehow mixed), and some physicists suggest space has actually more than 3 dimensions. But also in classical physics it is useful to use larger vectors, such as the $6D$ vector including the position and velocity of a body, or the $6ND$ vector with position and velocities of N particles (point-like bodies), a vector that identifies the state of a physical system.

nD vectors are also very common, or better ubiquitous, in computer science. The mathematician's generalisation to

$$\mathbf{x} = \{x_1, \dots, x_i, \dots, x_n\}, \tag{8}$$

$$\mathbf{z} = c\mathbf{x} \Rightarrow z_i = cx_i, \tag{9}$$

$$\mathbf{z} = c\mathbf{x} + \mathbf{y} \Rightarrow y_i = x_i + y_i, \tag{10}$$

is thus welcome.

3 Matrices: what and why

When we are considering points on a single line, we can easily transform between them just by multiplying by a constant a

$$y = ax. \tag{11}$$

Of course we may define more complex functions

$$y = f(x), \tag{12}$$

but the multiplication by a constant has some useful properties, namely that it respects sums and ratios between number since obviously

$$a(x + z) = ax + ay, \qquad a(cx) = cax \tag{13}$$

(this would not be true, for example, for $f(x) = x^2$).

How can we generalise this to the 3D space? And why do we need it?

Let us imagine we have the map of a room, with a given orientation and scale, and that we want to change both scale and orientation. Basically, we want to rotate and scale our map. But we still want our map to be reliable with respect with the "walking in this and that direction" discussion that we had before. Namely, if to vectors $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ and $\mathbf{w}$ correspond the mapped vectors $\mathbf{x}'$, $\mathbf{y}'$, $\mathbf{z}'$ and $\mathbf{w}'$, we want to have

$$\mathbf{z}' = \mathbf{x}' + \mathbf{y}' \;\text{ if }\; \mathbf{z} = \mathbf{x} + \mathbf{y} \tag{14}$$

and

$$\mathbf{w}' = c\,\mathbf{x}' \ \text{ if } \ \mathbf{w} = c\,\mathbf{x}. \tag{15}$$

We want thus to define a linear mapping $A\mathbf{x}$ such that

$$A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y}, \tag{16}$$

$$A(c\,\mathbf{y}) = c\,A\mathbf{x}. \tag{17}$$

Such a map is given by a matrix, a 3 by 3 table of numbers.

3.0.1 2D rotations

Let us start with a trivial 2D example. We have the vector

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{18}$$

If we rotate it counter clock-wise of 90 degrees, we obtain

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{19}$$

Let us write the matrix

$$R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{20}$$

and define the transformation (rotation)

$$\mathbf{y} = R\mathbf{x} = \begin{pmatrix} 0x_1 - x_2 \\ 1x_1 + 0x_2 \end{pmatrix}, \tag{21}$$

i.e. by "row by column" multiplication, we see that we obtain the wanted rotated vector.

In general, we have

$$\mathbf{y} = R\mathbf{x} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}. \tag{22}$$

You may check geometrically that this gives the expected rotation. For example

$$R\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}. \tag{23}$$

We have $\sin(-\pi/2) = -1$, $\cos(-\pi/2) = 0$. In general, we name

$$R(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{24}$$

a clock-wise rotation of an angle θ. Again you may check that

$$\mathbf{y} = R(\theta)\mathbf{x} = \begin{pmatrix} \cos\theta x_1 + \sin\theta x_2 \\ -\sin\theta x_1 + \cos\theta x_2 \end{pmatrix} \tag{25}$$

has the correct geometrical meaning.

3.0.2 Linearity

Are this application linear, i.e. do they satisfy eqs. (16,17)? The matrix multiplication rule assures that. Let us define a mapping from $\mathbf{x}$ to $\mathbf{x}'$ given by

$$x'_1 = \alpha x_1 + \beta x_2, \tag{26}$$

$$x'_2 = \gamma x_1 + \delta x_2. \tag{27}$$

It is straightforward to show that

$$z'_1 = \alpha(z_1) + b(z_2) = \alpha(x_1 + y_1) + \beta(x_2 + y_2) = (\alpha x_1 + \beta x_2) + (\alpha y_1 + \beta y_2) = x'_1 + y'_1, \tag{28}$$

$$z'_2 = \gamma(z_1) + \delta(z_2) = \gamma(x_1 + y_1) + \delta(x_2 + y_2) = (\gamma x_1 + \delta x_2) + (\gamma y_1 + \delta y_2) = x'_2 + y'_2, \tag{29}$$

and

$$w'_1 = \alpha(w_1) + \beta(w_2) = \alpha(cx_1) + \beta(cx_2) = c(\alpha x_1) + c(\beta x_2) = cx'_1, \tag{30}$$

$$w'_2 = \gamma(w_1) + \delta(w_2) = \gamma(cx_1) + \delta(cx_2) = c(\gamma x_1) + c(\delta x_2) = cx'_2. \tag{31}$$

For large n we will easily finish Latin and Greek letters, it is easier to write $\mathbf{y} = A\mathbf{x}$ where A is given by the 4 components

$$y_1 = a_{1,1}x_1 + a_{1,2}x_2, \tag{32}$$

$$y_2 = a_{2,1}x_1 + a_{2,2}x_2, \tag{33}$$

or in the general nD case

$$y_j = \sum_i a_{j,i}x_i, \tag{34}$$

which is the mathematical expression for the row by column rule.

3.0.3 Other examples

The matrix

$$\begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{35}$$

rotates on the x, y plane leaving the z direction unchanged. What are the corresponding rotations in the x, z and y, z planes?

The matrix

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \tag{36}$$

scales everything of a factor 2, while

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \tag{37}$$

scales only in the x direction (check it).

3.1 Matrix product

Originally we wanted to *rotate and scale* a map. Let us call

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{38}$$

the 90 degrees rotation matrix, and

$$B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \tag{39}$$

the scale 2 matrix. If we take

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{40}$$

and compute $\mathbf{y} = A\mathbf{x}$ we have

$$\mathbf{y} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}. \tag{41}$$

Now we may apply $\mathbf{z} = B\mathbf{y}$ and obtain

$$\mathbf{z} = \begin{pmatrix} -2x_2 \\ 2x_1 \end{pmatrix}. \tag{42}$$

This is equivalent to applying a

$$C = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix} \tag{43}$$

matrix directly to $\mathbf{x}$. This matrix is given by the matrix product rule

$$C\mathbf{z} = BA\mathbf{x} \Rightarrow C = BA. \tag{44}$$

In general, if

$$y_1 = a_{1,1}x_1 + a_{1,2}x_2, \tag{45}$$

$$y_2 = a_{2,1}x_1 + a_{2,2}x_2, \tag{46}$$

$$z_1 = b_{1,1}y_1 + b_{1,2}y_2, \tag{47}$$

$$z_2 = b_{2,1}y_1 + b_{2,2}y_2, \tag{48}$$

and

$$z_1 = c_{1,1}x_1 + c_{1,2}x_2, \tag{49}$$

$$z_2 = c_{2,1}x_1 + c_{2,2}x_2, \tag{50}$$

by substitution and commutative, associative, distributive properties we have

$$z_1 = (b_{1,1}a_{1,1} + b_{1,2}a_{2,1})x_1 + (b_{1,1}a_{1,2} + b_{1,2}a_{2,2})x_2, \tag{51}$$

$$z_2 = (b_{2,1}a_{1,1} + b_{2,2}a_{2,1})x_1 + (b_{2,1}a_{1,2} + b_{2,2}a_{2,2})x_2, \tag{52}$$

or

$$c_{1,1} = b_{1,1}a_{1,1} + b_{1,2}a_{2,1}, \tag{53}$$

$$c_{1,2} = b_{1,1}a_{1,2} + b_{1,2}a_{2,2}, \tag{54}$$

$$c_{2,1} = b_{2,1}a_{1,1} + b_{2,2}a_{2,1}, \tag{55}$$

$$c_{2,2} = b_{2,1}a_{1,2} + b_{2,2}a_{2,2}. \tag{56}$$

This generalises to the nD "row by column" rule

$$c_{j,i} = \sum_k b_{j,k}a_{k,i}. \tag{57}$$

3.2 Matrices do not commute

Consider a composite rotation, first of 90 degrees around the z axis, and then 90 degrees around the x axis

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}. \tag{58}$$

Let us change the order of rotations

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}. \tag{59}$$

The result is different! Is that a problem with our definition of matrix product? Try to understand it from a geometrical point of view, there is actually no problem. Rotations do not commute, the order matters!

You may better understand from the geometrical analysis of this simpler problem (apply to some simple vector)

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix}, \tag{60}$$

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 2 & 0 \end{pmatrix}. \tag{61}$$

Some trivial exercises.

Does the order of matrices matter in the product BA of eqs. (38,39)?

Can you show by matrix product that for 2D matrices (eq. 24)

$$R(\theta_1)R(\theta_2) = R(\theta_2)R(\theta_1) = R(\theta_1 + \theta_2) \tag{62}$$

applies? (Check the equation for the sine and cosine of summed angles).