Chapter 1

Preview: what is linear algebra about?

1.1 The setup: vector spaces

This module is an introduction to the theory, methods, practices and applications of linear algebra. Like many other areas of mathematics, linear algebra is concerned with a wide class of sets that have some particular structure or properties (these are basically the environments in which we work), and with some special functions between these sets, that have some sort of good behaviour with respect to the structure or properties that we are interested in. This is a fairly vague description of a system of organisation of ideas that is very prominent throughout mathematics, and can be expressed in more detail within certain frameworks. For example, in calculus, we are generally interested in functions from (subsets of) $\mathbb{R}$ to $\mathbb{R}$ - the set of all such things is more or less the environment that we work in. We are not really interested in *all* such functions, but maybe in the continuous functions or the differentiable ones, basically those functions that are amenable to study via the objects and themes of calculus.

In linear algebra, the environments that we work in are called *vector spaces*, and the main (but not only) functions of interest are called *linear transformations*. The entire setup is intricately bound up with the algebra of *matrices*. The subject and its methods have extraordinary prevalence, importance and applicability in every area of the mathematical sciences. For example, linear algebra allows for a translation of many problems of geometry into a concrete or computational setting, thanks to the innovation of coordinate geometry in or before the 17th century. Because linear algebra is relatively well understood and conducive to computer implementations, methods for solving complex problems that cannot be handled analytically often involve a reduction to or approximation by a problem of linear algebra. Linear algebra is central to the study of statistics, particularly in any situation where multiple random variables need to be considered simultaneously. A robust knowledge of the basic principles and methods of linear algebra, which we will study in this course, is essential for the study of virtually every area of mathematics and its applications (further examples include analysis, abstract algebra, combinatorics, theoretical physics and the theory of graphs and networks). Despite its position as part of the basic fabric of mathematics, linear algebra also continues to be a subject of research activity in its own right. NUI Galway will host the conference of the International Linear Algebra Society in 2022.

1.1.1 Vector Spaces

In order to define a vector space, we need to already have a system of scalars (or numbers) in mind. This set of scalars needs to have the structure of a *field*. This means that within the set of scalars, it is possible to add, subtract or multiply any two elements, and it is possible to divide any element by any non-zero element, *without moving outside the set*. [1] To get a sense of what a field is, we look at some examples.

[1] This is not a complete definition of a field. We also require that the addition and multiplication are both *commutative* and *associative*, and that they interact according to the *distributive laws*.

1. The set $\mathbb{R}$ of *real numbers* is a field.

2. The set $\mathbb{R}^+$ of *positive* real numbers is *not* a field. While it is possible to to add or multiply any pair of elements of $\mathbb{R}^+$ and stay in $\mathbb{R}^+$, the operation of subtracting one element of $\mathbb{R}^+$ from another can take us outside $\mathbb{R}^+$. For example 3 and 5 are elements of $\mathbb{R}^+$ but $3-5$ is not.

3. The set $\mathbb{Z}$ of integers is *not* a field. Adding, subtracting, or multiplying pairs of integers never takes us outside $\mathbb{Z}$. But attempting to divide one integer by another can take us outside $\mathbb{Z}$. For example 3 and 5 are elements of $\mathbb{Z}$ but $\frac{3}{5}$ is not.

4. The set $\mathbb{Q}$ of rational numbers *is* a field. It is defined by $\mathbb{Q} = \left\{ \frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0 \right\}$.

5. The set $\mathbb{C}$ of complex numbers *is* a field. Its elements are all expressions of the form $z = a + bi$, where a and b are real numbers (called the real and imaginary parts of the complex number z) and $i^2 = -1$.

For most students, it is likely that $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ are the most familiar and obvious examples of fields. Our "default" example will be the field $\mathbb{R}$ of real numbers, and it is fine to just think of the field as being $\mathbb{R}$ as we proceed, while you are getting used to the idea of a field. There are other examples however that will sometimes be of interest. One is the set $\mathbb{Z}/p\mathbb{Z}$ of integers modulo p, for any prime number p. We will write $\mathbb{F}_p$ for the integers modulo p. The elements of $\mathbb{F}_p$ are written as $0, 1, \dots, p-1$, with addition and multiplcation modulo p. For $p = 5$, the addition table and the multiplication table (for non-zero elements) are given below.

	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

Exercise 1.1.1. *If n is not prime, $\mathbb{Z}/n\mathbb{Z}$ is not a field. Why not? Try an example like $n = 6$ if you are not sure.*

Exercise 1.1.2. *Which of the following are fields? Explain!*

1. *The set of real numbers of the form $a + b\sqrt{2}$, where a and b are rational.*

2. *The set of complex numbers of the form $0 + bi$, where $b \in \mathbb{R}$.*

3. *The set of irrational real numbers.*

4. *The set of complex numbers of the form $a + bi$, where a and b are rational.*

Now we come to the most important definition of this course, that of a *vector space*. One thing to bear in mind is that our use of the term "vector" is not the same as the use of that term in physics to refer to a quantity that has both magnitude and direction, or its use to refer to a directed line segment or "arrow" in coordinate geometry. Our context is broader than either of those. Informally, a vector space V over a field $\mathbb{F}$ is a set whose elements can be added, subtracted and multiplied by elements of $\mathbb{F}$ (called *scalars* without ever moving outside the set V. Here are some examples.

1. The set $\mathbb{R}^3$ of *column vectors of length 3 over* $\mathbb{R}$ is a vector space over $\mathbb{R}$. Its elements are all vectors of the form $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$, where a, b, c are real numbers. They can be added, subtracted or multiplied by scalars as in these examples.

$$\begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix} + \begin{pmatrix} -3 \\ -5 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ -6 \\ 6 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ -5 \end{pmatrix} - \begin{pmatrix} -2 \\ 2 \\ -7 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \\ 2 \end{pmatrix}, \quad 4\begin{pmatrix} 0 \\ 6 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 24 \\ -8 \end{pmatrix}$$

2. The set $\mathbb{R}^{\mathbb{N}}$ of all infinite sequences of real numbers is a vector space over $\mathbb{R}$. To add two sequences together, we add their terms in every position. To subtract one sequence from another, we subtract its terms in every position. To multiply a sequence by a scalar k, we multiply every term by k.

3. The set of all 2×2 matrices with entries in $\mathbb{R}$ is a vector space over $\mathbb{R}$, with the usual addition and subtraction of matrices. To multiply a matrix by a scalar, we multiply each of its entries by that scalar.

4. $\mathbb{C}$ is a vector space over $\mathbb{R}$.

5. $\mathbb{R}$ is a vector space over $\mathbb{Q}$.

6. The set of all row vectors with three entries in $\mathbb{F}_2$ is a vector space over $\mathbb{F}_2$, with exactly eight elements. Here they are.

$$(0,0,0),\ (0,0,1),\ (0,1,0),\ (0,1,1),\ (1,0,0),\ (1,0,1),\ (1,1,0),\ (1,1,1).$$

In the first example above, we can observe that every element of $\mathbb{R}^3$ can be obtained by adding together a scalar multiple of $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, a scalar multiple of $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and a scalar multiple of $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. For example

$$\begin{pmatrix} 37 \\ \pi \\ -\frac{1}{5} \end{pmatrix} = 37 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \pi \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

We wil say that $\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$ is a *spanning set* of $\mathbb{R}^3$ (over $\mathbb{R}$). There is no finite subset of $\mathbb{R}^{\mathbb{N}}$ that is a spanning set of $\mathbb{R}^{\mathbb{N}}$ over $\mathbb{R}$. We say that $\mathbb{R}^{\mathbb{N}}$ is *infinite dimensional* as a vector space over $\mathbb{R}$.

Exercise 1.1.3. *This is something to think about, we will be better equipped shortly to answer it fully.*

1. *Prove the assertion above, that there is no finite spanning set of $\mathbb{R}^{\mathbb{N}}$ over $\mathbb{R}$ (this is tricky).*

2. *Give an example of a finite spanning set of $\mathbb{C}$ over $\mathbb{R}$.*

3. *Does $\mathbb{R}$ have a finite spanning set as a vector space over $\mathbb{Q}$? (also tricky at this point).*

For practical purposes, it is generally fine to think of a vector space (over a field $\mathbb{F}$) as a set whose elements can be added, subtracted, or multiplied by scalars. But this is not really adequate as a formal definition, we need to specify some details for example about how addition interacts with mutiplication by scalars. Here is the full definition.

Definition 1.1.4. *A vector space V over a field $\mathbb{F}$ is a non-empty set equipped with an addition operation $(+)$, and whose elements can be multiplied by scalars in $\mathbb{F}$, subject to the following axioms.*

1. *For all $u, v \in V$, $u + v = v + u$ (addition is commutative).*

2. *For all $u, v, w \in V$, $(u + v) + w = u + (v + w)$ (addition is associative).*

3. *V includes an element 0_V, with the property that $0_V + v = v$ for all $v \in V$ (zero element of V or zero vector).*

4. *For every $v \in V$, there exists an element $-v$ of V, with the property that $v + (-v) = 0_V$ (subtraction).*

5. *If $\alpha, \beta \in \mathbb{F}$ and $v \in V$, then $\alpha(\beta v) = \alpha\beta(v)$ (compatibility of scalar multiplication with multiplication in $\mathbb{F}$).*

6. *If $\alpha \in \mathbb{F}$ and $u, v \in V$, then $\alpha(u + v) = \alpha u + \alpha v$ (distributivity of scalar mulatiplication over addition in V).*

7. *If $\alpha, \beta \in \mathbb{F}$ and $v \in V$, then $(\alpha + \beta)v = \alpha v + \beta v$ (distributivity of scalar mulatiplication over addition in $\mathbb{F}$).*

8. *$1_{\mathbb{F}} v = v$ for all $v \in V$, where $1_{\mathbb{F}}$ is the multplicative identity element of $\mathbb{F}$.*

When working with examples, we rarely have to worry very much about any of the technical specifications in Definition 1.1.4; they tend to be clearly satisfied. But there is a value in devoting some time to studying each of these axioms carefully. When a class of objects (like vector spaces) is defined in terms of the algebraic operations, it is important to specify the properties that the operations must have in order to qualify for admission to the class. Even if these propoerties are natural and obvious in the examples that motivate the definition, it is still important to note them systematically, they form the rules of the game and a key reference point in the development of the theory.

Exercise 1.1.5. *Check that the axioms are satisfied in the examples that we have met so far.*

1.2 Review of Matrix Algebra

A $m \times n$ matrix over a field $\mathbb{F}$ is an an array of m rows and n columns, whose entries are elements of $\mathbb{F}$. We can take $\mathbb{F}$ to be the field of real numbers. The expression $m \times n$ is referred to as the *size* of a matrix (even though what it really describes is the *shape*). Two matrices can be added together if they have the same size; in this case their sum is obtained by just adding the entries in each position. The $m \times n$ zero matrix is the $m \times n$ matrix whose entries are all zeros. It is the *identity element* for addition of $m \times n$ matrices - this means that addition it to another $m \times n$ matrix has no effect. A matrix can be multiplied by a scalar; this means multiplying each of its entries by that scalar. With these operations of addition and scalar multiplication, the set of $m \times n$ matrices over a field $\mathbb{F}$ is a *vector space* over $\mathbb{F}$.
Notation: We use the notation $M_{m \times n}(\mathbb{F})$ for the vector space of all $m \times n$ matrices over $\mathbb{F}$. When $m = n$, we abbreviate this to $M_n(\mathbb{F})$.

Example 1.2.1. *In* $M_{2\times 3}(\mathbb{R})$,

$$2\begin{pmatrix} 1 & 0 & -1 \\ 2 & -5 & 1 \end{pmatrix} - 3\begin{pmatrix} 2 & 4 & -1 \\ 0 & 1 & -3 \end{pmatrix} = \begin{pmatrix} 2(1)-3(2) & 2(0)-3(4) & 2(-1)-3(-1) \\ 2(2)-3(0) & 2(-5)-3(1) & 2(-3)-3(-3) \end{pmatrix} = \begin{pmatrix} -4 & -12 & 1 \\ 4 & -13 & 3 \end{pmatrix}.$$

We can sometimes also *multiply* matrices, but the way to do this is not as obvious. We begin with a few definitions.

Definition 1.2.2. *Suppose that $v_1, v_2, \dots, v_k$ are elements of a vector space V over a field $\mathbb{F}$. A $\mathbb{F}$-linear combination (or just linear combination) of $v_1, \dots, v_k$ is an element of V that has the form $a_1v_1 + a_2v_2 + \cdots + a_kv_k$, wher the a_i are elements of $\mathbb{F}$. In this situation the a_i are called the* coefficients *in the linear combination.*

The example above shows a linear combination of two matrices in $M_{2\times 3}(\mathbb{R})$, with coefficients 2 and -3.

Definition 1.2.3. *A* column vector *is a matrix with one column. A* row vector *is a matrix with one row.*

Before defining matrix multiplication in general, we define the product of a matrix with a column vector (when that exists).

Definition 1.2.4. *Let A be a $m \times n$ matrix and let v be a column vector with n entries. Then the matrix-vector product Av is the column vector obtained by taking the linear combination of the columns of A whose coefficients are the entries of v. It is a column vector with m entries.*

Example 1.2.5. $\begin{pmatrix} 2 & 4 & -1 \\ 0 & 1 & -3 \end{pmatrix}\begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} = 3\begin{pmatrix} 2 \\ 0 \end{pmatrix} + (-2)\begin{pmatrix} 4 \\ 1 \end{pmatrix} + 1\begin{pmatrix} -1 \\ -3 \end{pmatrix} = \begin{pmatrix} -3 \\ -5 \end{pmatrix}$