Notes 13 – Vector Spaces

The theory of linear combinations, linear independence, bases, and subspaces that we have studied in relation to $\mathbb{R}^n$ can be generalized to the more general study of *vector spaces.* Any subspace of $\mathbb{R}^n$ (including of course $\mathbb{R}^n$ itself) is an example of a vector space, but there are many others including sets of matrices, polynomials and functions.

L13.1 Motivation. A subspace of $\mathbb{R}^n$ is the prime example of a *vector space,* but there are a number of reasons for discussing the general definition, namely

(i) to emphasize aspects of the theory that do *not* depend upon the choice of a specific basis,

(ii) to allow the use of scalars that are different from real numbers,

(iii) to extend the theory to function spaces of infinite dimensions.

We shall explain each of these points in turn.

(i) The whole description of $\mathbb{R}^n$ is modelled on the existence of its canonical basis. To be specific, consider $\mathbb{R}^{n,1}$ and let $\mathbf{e}_j$ denote the jth column of the identity matrix I_n. Then a typical element of $\mathbb{R}^{n,1}$ is given by

$$\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ x_n \end{pmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n,$$

and is represented by its coefficients relative to the basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$. But when we wish to describe subspaces of $\mathbb{R}^{n,1}$ there is a need work with other bases. In fact, any subspace of $\mathbb{R}^n$ is a vector space in its own right. In general it is important to be able *change* basis; in this way, the abstract concept of *vector space* comes into its own.

(ii) The 'scalars' that are used to multiply vectors in the definition of a vector space need not be real numbers. The set of scalars is required to be what is called a *field,* of which $\mathbb{R}$ is only one example. Other examples of fields include the set $\mathbb{Q}$ of rational numbers (p/q where p, q are integers with $q \neq 0$), the set $\mathbb{C}$ of complex numbers ($x+iy$ with $x, y \in \mathbb{R}$ and $i = \sqrt{-1}$), and the 'binary set' $B = \{0, 1\}$ (also called $\mathbb{F}_2$) consisting of just two elements.

(iii) In this course, we shall only work with vector spaces of *finite* dimension. We shall explain that such a vector space V is characterized by the existence of a basis of finite size n. The choice of such a basis makes V closely resemble the set $\mathbb{R}^n$ or (for the other choices of fields mentioned above) $\mathbb{Q}^n$, $\mathbb{C}^n$ or the finite set B^n of size 2^n. However, in analysis, the most important examples of vector spaces do not fall into this category and once again ones needs to rely upon the abstract and basis-independent theory.

L13.2 The definition of a vector space. In order to define a vector space in general, one first needs a *field* F of scalars. For the moment, we shall suppose that F is one of $\mathbb{R}, \mathbb{Q}, \mathbb{C}, B$. The important thing about F one needs to know is that its elements can be added, subtracted, multiplied and divided, and that there are two special ones, 0 and 1.

A vector space is a set V in which it is possible to form

(i) the *sum* $\mathbf{u} + \mathbf{v}$ of $\mathbf{u}, \mathbf{v} \in V$,

(ii) the *product* $a\mathbf{v}$ of $a \in F$ with $\mathbf{v} \in V$.

The elements of V are called 'vectors' even though they do not necessarily resemble vectors in $\mathbb{R}^n$. The two basic operations are subject to a number of rules that formalize the ones that are completely obvious in the case $F = \mathbb{R}$ and $V = \mathbb{R}^n$ we are most familar with. There is no need to memorize these rules, as they are quickly absorbed in practice:

Definition. *V is said to be a* vector space over *F provided*
(a) addition of vectors behaves like addition of real numbers in that it satisfies

$$\begin{aligned} &(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w}), \\ &\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}, \qquad\qquad \textit{for all } \mathbf{u},\mathbf{v},\mathbf{w} \in V, \end{aligned} \tag{1}$$

there is a zero *or* null *element $\mathbf{0}$ for which*

$$\mathbf{0}+\mathbf{v} = \mathbf{v} \qquad \textit{for all } \mathbf{v} \in V,$$

and each vector $\mathbf{v} \in V$ has a 'negative' $-\mathbf{v}$ with the property that $\mathbf{v}+(-\mathbf{v}) = \mathbf{0}$;
(b) the 'internal' operations of F are compatible with (i) and (ii) in the sense that

$$\begin{aligned} (a+b)\mathbf{v} &= a\mathbf{v}+b\mathbf{v} \\ (ab)\mathbf{v} &= a(b\mathbf{v}), \\ a(\mathbf{u}+\mathbf{v}) &= a\mathbf{u}+a\mathbf{v}, \end{aligned} \tag{2}$$

and finally,

$$1\mathbf{v} = \mathbf{v}.$$

On this page, we have been careful to type elements of V (but not F) in boldface, though in handwriting one does not normally distinguish elements of V in any way. It is important to observe that instances of both addition and multiplication in (2) occur with different meanings. The conditions in (a), taken together, assert that the operation + makes V into what is called a *commutative* (or *abelian) group.* Of course, we write $\mathbf{u}+(-\mathbf{v})$ as $\mathbf{u}-\mathbf{v}$, and this process defines *subtraction* in a vector space.

Here is a simple consequence of the axioms above:

$$0\mathbf{v}+0\mathbf{v} = (0+0)\mathbf{v} = 0\mathbf{v}.$$

The various rules in (a) allow us to subtract $0\mathbf{v}$ (without knowing what this equals) to get

$$0\mathbf{v} = \mathbf{0},$$

so that $0\mathbf{v}$ is always the null vector.

Example. To keep matters familiar, we first suppose that $F = \mathbb{R}$, in which case V is called a *real* vector space. Certainly $V = \mathbb{R}^n$ satisfies the definition above with the usual operations that we have used repeatedly.

But another example is to take V to be the set $\mathbb{R}^{m,n}$ of matrices of size $m \times n$. We explained in the first lecture how to add such matrices together, and multiply them by scalars. From the point of view of vector spaces (in which multiplication of matrices plays no part), there is little difference between $\mathbb{R}^{m,n}$ and the space (of say row vectors) $\mathbb{R}^{1,mn}$. For example, we can pass from $\mathbb{R}^{2,3}$ to $\mathbb{R}^6$ by the correspondence

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \leftrightarrow (a,b,c,d,e,f).$$

it does not matter whether we use the left-hand or right-hand description to define the two basic operations – the result is the same. But we could equally well have chosen to represent the matrix with (c,f,b,e,a,d); for this reason the vector spaces $\mathbb{R}^{m,n}$, $\mathbb{R}^{1,mn}$ are not *identical.*

L13.3 Polynomials and functions. A more original example is obtained using polynomials. Recall that a *polynomial* is an expression of the form

$$\mathbf{p}(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n. \tag{3}$$

We are most familiar with the case in which the coefficients are real numbers, but they could belong to a field F. The polynomial has *degree* equal to n provided $a_n \neq 0$. We have written $\mathbf{p}$ in boldface to emphasize that is is to be treated as a 'vector', though it is also the *function*

$$x \mapsto \mathbf{p}(x);$$

the choice of symbol for the variable is irrelevant, and one often writes $\mathbf{p}(t)$.

The *constant term*

$$a_0 = \mathbf{p}(0)$$

of the polynomial is none other than the value of the function at 0. It vanishes if and only if the polynomial $\mathbf{p}(x)$ is *divisible* by x.

Proposition. *The set $F_n[x]$ of polynomials (in a variable x) of degree no more than n with coefficients in a field F is a vector space over F.*

Proof. There is way to define define the basic operations, using a rule that works for any functions. Namely, we set

$$\begin{aligned} (\mathbf{p}_1 + \mathbf{p}_2)(x) &= \mathbf{p}_1(x) + \mathbf{p}_2(x), \\ (a\,\mathbf{p})(x) &= a\,\mathbf{p}(x), \qquad a \in F. \end{aligned} \tag{4}$$

Using (3), it is obvious that the sum of two polynomials is a polynomial, and that the product of a polynomial with a scalar is a polynomial. In practice, it is just a matter of applying the operations coefficient-wise, as in the example

$$(1+x)^2 + 3(1 + x + \tfrac{1}{2}x^2 + \tfrac{1}{6}x^3) = 4 + 5x + \tfrac{5}{2}x^2 + \tfrac{1}{2}x^3,$$

representing a LC of two elements in $\mathbb{R}_3[x]$. QED

The previous proposition is a special case of the

Proposition. *Let V be a vector space, and let $\mathscr{A}$ be* any *non-empty set. Then the rules (4) make the set of all mappings $f\colon \mathscr{A} \to V$ into a vector space.*

L13.4 More about fields.

We shall not give the formal definition of a field. But it is a set F that satisfies the rules of a vector space, in which we are allowed to take the set of scalars to be the same set F. Multiplication between scalars and vectors therefore becomes a multiplication between elements of F such that

$$1a = a, \qquad a \in F,$$

and 1 is the *multiplicative identity* or *unit.* The multiplication is required to be commutative, so that

$$ab = ba, \qquad a, b \in F, \tag{5}$$

and every nonzero element $a \in F$ must have a *multiplicative inverse,* written a^{-1}, satisfying

$$a\, a^{-1} = 1.$$

Every field must contain at least two elements: the additive identity (usually written 0) and the multiplicative identity (written 1). If there are no other elements, we obtain $B = \{0, 1\}$ that *is* a field with the operations

$$\begin{gathered} 0+0=0, \quad 0+1=1=1+0, \quad 1+1=0, \\ 0\,0=0, \quad 0\,1=0=1\,0, \quad 1\,1=1. \end{gathered} \tag{6}$$

Example. Here is an example of a field F with 4 elements. It will be defined as a vector space over a simpler field, namely B. The set F consists of all linear combinations

$$b_1\mathbf{f}_1 + b_2\mathbf{f}_2, \qquad b_1, b_1 \in B,$$

in which we decree that $\mathbf{f}_1, \mathbf{f}_2$ are independent. Although b_1, b_2 are arbitrary, there are only two choices for each. We can therefore list all four elements of F as row vectors

$$\begin{aligned} (0,0) &= 0\mathbf{f}_1 + 0\mathbf{f}_2 = 0, \\ (1,0) &= 1\mathbf{f}_1 + 0\mathbf{f}_2 = f_1, \\ (0,1) &= 0\mathbf{f}_1 + 1\mathbf{f}_2 = f_2, \\ (1,1) &= 1\mathbf{f}_1 + 1\mathbf{f}_2 = 1. \end{aligned}$$

(On the right, we have avoided boldface to emphazise that the elements are to be treated like numbers, not vectors.) Multiplication is carried out component-wise, using the operations of B. The reason for also calling the last element 1 is that $(1,1)(a,b) = (1a, 1b) = (a,b)$ for $a, b \in B$. The full multiplication table for F is symmetric because of (5):

$\cdot$	0	f_1	f_2	1
0	0	0	0	0
f_1	0	f_1	1	f_1
f_2	0	1	f_2	f_2
1	0	f_1	f_2	1

If p is a prime number, the set $\{0, 1, 2, \ldots, p-1\}$ with addition and multiplication modulo p ('clockface arithmetic') becomes a field with exactly p elements. Applying the construction of the Example with $\mathbf{f}_1, \ldots, \mathbf{f}_k$ in place of $\mathbf{f}_1, \mathbf{f}_2$ shows that there is a field with p^k elements for any positive integer k. It turns out that this is essentially the *only* field with p^k elements. Moreover, *any* finite field has p^k elements for some prime number $p \geqslant 2$ and integer $k \geqslant 1$.

L13.5 Further exercises.

1. Show that the set of all *differentiable* functions $f\colon (0,1) \to \mathbb{R}$ is a real vector space. By considering polynomials, or otherwise, show that it is not finite-dimensional.

2. Let $V = B^{2,2}$ be the set of 2×2 matrices, each of whose entries is 0 or 1.
(i) How many elements does V have?
(ii) Show that V is a vector space with field B.
(iii) How many matrices in V have determinant (calculated using (6)) equal to 1?