PART I

Preliminaries

Cambridge University Press
978-0-521-89964-2 — An Introduction to Functional Analysis
James C. Robinson
Excerpt
More Information

1
Vector Spaces and Bases

Much of the theory of 'functional analysis' that we will consider in this book is an infinite-dimensional version of results familiar for linear operators between finite-dimensional vector spaces. We therefore start by recalling some of the basic theory of linear algebra, beginning with the formal definition of a vector space. We then discuss linear maps between vector spaces, and end by proving that every vector space has a basis using Zorn's Lemma. Proofs of basic results from linear algebra can be found in Friedberg et al. (2004) or in Chapter 4 of Naylor and Sell (1982), for example.

1.1 Definition of a Vector Space

The linear spaces that occur naturally in functional analysis are vector spaces defined over $\mathbb{R}$ or $\mathbb{C}$; we will refer to real or complex vector spaces respectively, but generally we will omit the word 'real' or 'complex' unless we need to make an explicit distinction between the two cases.

Throughout the book we use the symbol $\mathbb{K}$ to denote either $\mathbb{R}$ or $\mathbb{C}$.

Definition 1.1 A *vector space* V *over* $\mathbb{K}$ is a set V along with notions of addition in V and multiplication by scalars, i.e.

$$x + y \in V \quad \text{for} \quad x, y \in V \qquad \text{and} \qquad \lambda x \in V \quad \text{for} \quad \lambda \in \mathbb{K},\ x \in V, \tag{1.1}$$

such that

(i) additive and multiplicative identities exist: there exists a zero element $0 \in V$ such that $x + 0 = x$ for all $x \in V$; and $1 \in \mathbb{K}$ is the identity for scalar multiplication, $1x = x$ for all $x \in V$;
(ii) there are additive inverses: for every $x \in V$ there exists an element $-x \in V$ such that $x + (-x) = 0$;

(iii) addition is commutative and associative,

$$x + y = y + x \qquad \text{and} \qquad x + (y + z) = (x + y) + z,$$

for all $x, y, z \in V$; and

(iv) multiplication is associative,

$$\alpha(\beta x) = (\alpha\beta)x \qquad \text{for all} \qquad \alpha, \beta \in \mathbb{K},\ x \in V,$$

and distributive,

$$\alpha(x + y) = \alpha x + \alpha y \qquad \text{and} \qquad (\alpha + \beta)x = \alpha x + \beta x$$

for all $\alpha, \beta \in \mathbb{K}$, $x, y \in V$.

In checking that a particular collection V is a vector space over $\mathbb{K}$, properties (i)–(iv) are often immediate; one usually has to check only that V is closed under addition and scalar multiplication (i.e. that (1.1) holds).

1.2 Examples of Vector Spaces

Of course, $\mathbb{R}^n$ is a real vector space over $\mathbb{R}$; but is not a vector space over $\mathbb{C}$, since $\mathrm{i}\boldsymbol{x} \notin \mathbb{R}^n$ for any[1] $\boldsymbol{x} \in \mathbb{R}^n$. In contrast, $\mathbb{C}^n$ can be a vector space over both $\mathbb{R}$ and $\mathbb{C}$; the space $\mathbb{C}^n$ over $\mathbb{R}$ is (according to the terminology introduced above) a 'real vector space'. This example is a useful illustration that the real/complex label refers to the field $\mathbb{K}$, i.e. the allowable scalar multiples, rather than to the elements of the space itself.

Given any two vector spaces V_1 and V_2 over $\mathbb{K}$, the product space $V_1 \times V_2$ consisting of all pairs (v_1, v_2) with $v_1 \in V_1$ and $v_2 \in V_2$ is another vector space if we define

$$(v_1, v_2) + (u_1, u_2) := (v_1 + u_1, v_2 + u_2) \qquad \text{and} \qquad \alpha(v_1, v_2) := (\alpha v_1, \alpha v_2),$$

for $v_1, u_1 \in V_1$, $v_2, u_2 \in V_2$, $\alpha \in \mathbb{K}$.

We now introduce some less trivial examples.

Example 1.2 The space $\mathcal{F}(U, V)$ of all functions $f : U \to V$, where U and V are both vector spaces over the same field $\mathbb{K}$, is itself a vector space, if we use the obvious definitions of what addition and scalar multiplication should mean for functions. We give these definitions here for the one and only time:

[1] Throughout this book we will use a bold $\boldsymbol{x}$ for elements of $\mathbb{R}^n$ (also of $\mathbb{C}^n$), with $\boldsymbol{x}$ given in components by $\boldsymbol{x} = (x_1, \ldots, x_n)$.

for $f, g \in \mathcal{F}(U, V)$ and $\alpha \in \mathbb{K}$, we denote by $f + g$ the function from U to V whose values are given by

$$(f + g)(x) = f(x) + g(x), \qquad x \in U,$$

('pointwise addition') and by αf the function whose values are

$$(\alpha f)(x) = \alpha f(x), \qquad x \in U$$

('pointwise multiplication').

Example 1.3 The space $C([a, b]; \mathbb{K})$ of all $\mathbb{K}$-valued continuous functions on the interval $[a, b]$ is a vector space. We will often write $C([a, b])$ for $C([a, b]; \mathbb{R})$.

Proof The sum of two continuous functions is again continuous, as is any scalar multiple of a continuous function. □

Example 1.4 The space $\mathcal{P}(I)$ of all real polynomials on any interval $I \subset \mathbb{R}$,

$$\mathcal{P}(I) = \left\{ p\colon I \to \mathbb{R} : \ p(x) = \sum_{j=0}^{n} a_j x^j, \ n = 0, 1, 2, \ldots, \ a_j \in \mathbb{R} \right\}$$

is a vector space.

The next example introduces a family of spaces that will prove to be particularly important.

Example 1.5 For $1 \le p < \infty$ the space $\ell^p(\mathbb{K})$ consists of all pth power summable sequences $\boldsymbol{x} = (x_j)_{j=1}^{\infty}$ with elements in $\mathbb{K}$, i.e.

$$\ell^p(\mathbb{K}) = \left\{ \boldsymbol{x} = (x_j)_{j=1}^{\infty} : \ x_j \in \mathbb{K}, \ \sum_{j=1}^{\infty} |x_j|^p < \infty \right\}.$$

For $p = \infty$, $\ell^\infty(\mathbb{K})$ is the space of all bounded sequences in $\mathbb{K}$. Sometimes we will simply write ℓ^p for $\ell^p(\mathbb{K})$. Note that, as with $\mathbb{K}^n$, we will use a bold $\boldsymbol{x}$ to denote a particular sequence in ℓ^p.

For $\boldsymbol{x}, \boldsymbol{y} \in \ell^p(\mathbb{K})$ we set

$$\boldsymbol{x} + \boldsymbol{y} := (x_1 + y_1, x_2 + y_2, \ldots),$$

and for $\alpha \in \mathbb{K}$, $\boldsymbol{x} \in \ell^p$, we define

$$\alpha \boldsymbol{x} := (\alpha x_1, \alpha x_2, \ldots).$$

With these definitions $\ell^p(\mathbb{K})$ is a vector space.

Proof The only thing that is not immediate is whether $\boldsymbol{x} + \boldsymbol{y} \in \ell^p(\mathbb{K})$ if $\boldsymbol{x}, \boldsymbol{y} \in \ell^p(\mathbb{K})$. This is clear when $p = \infty$, since

$$\sup_{j\in\mathbb{N}} |x_j + y_j| \le \sup_{j\in\mathbb{N}} |x_j| + \sup_{j\in\mathbb{N}} |y_j| < \infty.$$

For $1 \le p < \infty$ this follows using the inequality

$$(a+b)^p \le [2\max(a,b)]^p \le 2^p(a^p + b^p), \qquad \text{for } a, b \ge 0; \tag{1.2}$$

for every $n \in \mathbb{N}$ we have

$$\sum_{j=1}^{n} |x_j + y_j|^p \le \sum_{j=1}^{n} 2^p(|x_j|^p + |y_j|^p) \le 2^p \sum_{j=1}^{\infty} |x_j|^p + 2^p \sum_{j=1}^{\infty} |y_j|^p < \infty$$

and so $\sum_{j=1}^{\infty} |x_j + y_j|^p < \infty$ as required. □

(The factor 2^p in (1.2) can be improved to 2^{p-1}; see Exercise 1.1.)

1.3 Linear Subspaces

If V is a vector space (over $\mathbb{K}$) then any subset $U \subset V$ is a *subspace* of V if U is again a vector space, i.e. if it is closed under addition and scalar multiplication, i.e. $u_1 + u_2 \in U$ for every $u_1, u_2 \in U$ and $\lambda u \in U$ for every $\lambda \in \mathbb{K}$, $u \in U$.

Example 1.6 For any $\boldsymbol{y} \in \mathbb{R}^n$, the set

$$\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} \cdot \boldsymbol{y} = 0\}$$

is a subspace of $\mathbb{R}^n$.

Example 1.7 The set

$$X = \left\{ f \in C([-1,1]) : \int_{-1}^{0} f(x)\,\mathrm{d}x = 0, \ \int_{0}^{1} f(x), \mathrm{d}x = 0 \right\}$$

is a subspace of $C([-1,1])$.

Example 1.8 The space $c_0(\mathbb{K})$ of all null sequences, i.e. of all sequences $\boldsymbol{x} = (x_j)_{j=1}^{\infty}$ such that $x_j \to 0$ as $j \to \infty$, is a subspace of $\ell^\infty(\mathbb{K})$, and for every $1 \le p < \infty$ the space $\ell^p(\mathbb{K})$ is a subspace of $c_0(\mathbb{K})$.

The space $c_{00}(\mathbb{K})$ of all sequences with only a finite number of non-zero terms is a subspace of $c_0(\mathbb{K})$ and of $\ell^p(\mathbb{K})$ for every $1 \le p \le \infty$.

Cambridge University Press
978-0-521-89964-2 — An Introduction to Functional Analysis
James C. Robinson
Excerpt
More Information

Proof For the inclusion properties of $c_0(\mathbb{K})$, note that any convergent sequence (in particular any null sequence) is bounded, which shows that $c_0(\mathbb{K}) \subset \ell^\infty(\mathbb{K})$. If $\boldsymbol{x} \in \ell^p$, $1 \le p < \infty$, then $\sum_{j=1}^\infty |x_j|^p < \infty$, which implies that $|x_j|^p \to 0$ as $j \to \infty$, so $\boldsymbol{x} \in c_0(\mathbb{K})$. The properties of $c_{00}(\mathbb{K})$ are immediate. □

1.4 Spanning Sets, Linear Independence, and Bases

We now recall the definition of a vector-space basis, which will also allow us to define the dimension of a vector space.

Definition 1.9 The *linear span* of a subset E of a vector space V is the collection of all finite linear combinations of elements of E:

$$\operatorname{Span}(E) = \left\{ v \in V : v = \sum_{j=1}^n \alpha_j e_j, \text{ for some } n \in \mathbb{N},\ \alpha_j \in \mathbb{K},\ e_j \in E \right\}.$$

We say that E *spans* V if $V = \operatorname{Span}(E)$.

If E spans V this means that we can write any $v \in V$ in the form

$$v = \sum_{j=1}^n \alpha_j e_j,$$

i.e v can be expressed as a finite linear combination of elements of E. (Once we have a way to discuss convergence we will also be able to consider 'infinite linear combinations', but these are not available when we can only use the vector-space axioms.)

Definition 1.10 A set $E \subset V$ is *linearly independent* if any finite collection of elements of E is linearly independent, i.e.

$$\sum_{j=1}^n \alpha_j e_j = 0 \qquad \Rightarrow \qquad \alpha_1 = \cdots = \alpha_n = 0$$

for any choice of $n \in \mathbb{N}$, $\alpha_j \in \mathbb{K}$, and $e_j \in E$.

To distinguish the standard definition of a basis for a vector space from the notion of a 'Schauder basis', which we will meet later, we refer to such a basis as a 'Hamel basis'.

Definition 1.11 A *Hamel basis* for a vector space V is any linearly independent spanning set.

Expansions in terms of basis elements are unique (for a proof see Exercise 1.3).

Lemma 1.12 *If E is a Hamel basis for V, then any element of V can be written uniquely in the form*

$$v = \sum_{j=1}^{n} \alpha_j e_j$$

for some $n \in \mathbb{N}$, $\alpha_j \in \mathbb{K}$, and $e_j \in E$.

Any Hamel basis E of V must be a maximal linearly independent set, i.e. E is linearly independent and $E \cup \{v\}$ is not linearly independent for any $v \in V \setminus E$. We now show that this can be reversed.

Lemma 1.13 *If $E \subset V$ is maximal linearly independent set, then E is a Hamel basis for V.*

Proof To show that E is a Hamel basis we only need to show that it spans V, since it is linearly independent by assumption.

If E does not span V, then there exists some $v \in V$ that cannot be written as any finite linear combination of the elements of E. To obtain a contradiction, we show that in this case $E \cup \{v\}$ must be a linearly independent set. Choose $n \in \mathbb{N}$ and $\{e_j\}_{j=1}^{n} \in E$, and suppose that

$$\sum_{j=1}^{n} \alpha_j e_j + \alpha_{n+1} v = 0.$$

Since v cannot be written as a sum of any finite collection of the $\{e_j\}$, we must have $\alpha_{n+1} = 0$, which leaves $\sum_{j=1}^{n} \alpha_j e_j = 0$. However, since E is linearly independent and $\{e_j\}_{j=1}^{n}$ is a finite subset of E it follows that $\alpha_j = 0$ for all $j = 1, \ldots, n$. Since we already have $\alpha_{n+1} = 0$, it follows that $E \cup \{v\}$ is linearly independent, contradicting the fact that E is a maximal linearly independent set. So E spans V, as claimed. □

If V has a basis consisting of a finite number of elements, then every basis of V contains the same number of elements (for a proof see Exercise 1.4).

Lemma 1.14 *If V has a basis consisting of n elements, then every basis for V has n elements.*

This result allows us to make the following definition of the dimension of a vector space.

Definition 1.15 If V has a basis consisting of a finite number of elements, then V is *finite-dimensional* and the *dimension* of V is the number of elements in this basis. If V has no finite basis, then V is *infinite-dimensional*.

Since a basis is a maximal linearly independent set (Lemma 1.13), it follows that a space is infinite-dimensional if and only if for every $n \in \mathbb{N}$ one can find a set of n linearly independent elements of V.

Example 1.16 For every $1 \le p \le \infty$ the space $\ell^p(\mathbb{K})$ is infinite-dimensional.

Proof Let us define for each $j \in \mathbb{N}$ the sequence

$$\boldsymbol{e}^{(j)} = (0, 0, \dots, 1, 0, \dots), \tag{1.3}$$

which consists entirely of zeros apart from having 1 as its jth term. We can also write

$$e_i^{(j)} = \delta_{ij} := \begin{cases} 1 & i = j \\ 0 & i \neq j, \end{cases} \tag{1.4}$$

where δ_{ij} is the Kronecker delta. These are all elements of $\ell^p(\mathbb{K})$ for every $p \in [1, \infty]$, and will frequently prove useful in what follows.

For any $n \in \mathbb{N}$ the n elements $\{\boldsymbol{e}^{(j)}\}_{j=1}^n$ are linearly independent, since

$$\sum_{j=1}^n \alpha_j \boldsymbol{e}^{(j)} = (\alpha_1, \alpha_2, \dots, \alpha_n, 0, 0, 0, \dots) = \boldsymbol{0}$$

implies that $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. It follows that $\ell^p(\mathbb{K})$ is an infinite-dimensional vector space. □

Example 1.17 The vector space $C([0, 1]; \mathbb{K})$ is infinite-dimensional.

Proof For any $n \in \mathbb{N}$ the functions $\{1, x, x^2, \dots, x^n\}$ are linearly independent: if

$$f(x) := \sum_{j=0}^n \alpha_j x^j = 0 \qquad \text{for every } x \in [0, 1],$$

then $\alpha_j = 0$ for every j. To see this, first set $x = 0$, which shows that $\alpha_0 = 0$, then differentiate once to obtain

$$f'(x) = \sum_{j=1}^{n} \alpha_j j x^{j-1} = 0$$

and set $x = 0$ to show that $\alpha_1 = 0$. Continue differentiating repeatedly, each time setting $x = 0$ to show that $\alpha_j = 0$ for all $j = 0, \ldots, n$. □

1.5 Linear Maps between Vector Spaces and Their Inverses

Vector spaces have a linear structure, i.e. we can add elements and multiply by scalars. When we consider maps from one vector space to another, it is natural to consider maps that respect this linear structure.

Definition 1.18 If X and Y are vector spaces over $\mathbb{K}$, then a map $T : X \to Y$ is *linear* if

$$T(x + x') = T(x) + T(x') \quad \text{and} \quad T(\alpha x) = \alpha T(x), \qquad \alpha \in \mathbb{K}, \ x, x' \in X.$$

(This is the same as requiring that $T(\alpha x + \beta x') = \alpha T(x) + \beta T(x')$ for any $\alpha, \beta \in \mathbb{K}$, $x, x' \in U$.)

We often omit the brackets around the argument, and write Tx for $T(x)$ when T is linear.

Note that the definition of what it means to be linear involves the field $\mathbb{K}$. So, for example, if we take $X = Y = \mathbb{C}$ and let $T(z) = \overline{z}$ (the complex conjugate of z), this map is linear if we take $\mathbb{K} = \mathbb{R}$, but not if we take $\mathbb{K} = \mathbb{C}$. We always have

$$T(z + w) = \overline{z + w} = \overline{z} + \overline{w} = T(z) + T(w), \qquad z, w \in \mathbb{C},$$

but the linearity property for scalar multiples only holds if $\alpha \in \mathbb{R}$, since

$$T(\alpha z) = \overline{\alpha z} = \overline{\alpha}\,\overline{z}$$

and this is equal to $\alpha\overline{z} = \alpha T(z)$ if and only if $\alpha \in \mathbb{R}$.

This kind of ‘conjugate-linear’ behaviour is common enough that it is worth making a formal definition.