Linear Functional Analysis

线性泛函分析

Bryan P. Rynne
Martin A. Youngson

Tsinghua University Press   Springer

Linear Functional Analysis

线性泛函分析

Bryan P. Rynne
Martin A. Youngson

Tsinghua University Press
Beijing

Springer

内容提要

本书以较小的篇幅介绍了线性泛函分析的基本内容：赋范空间和 Banach 空间，内积空间和 Hilbert 空间，线性算子，紧算子及其在积分方程和微分方程中的应用。本书内容深入浅出、通俗易懂，重要的概念和定理均有背景介绍，并配有简单例子加以解释；排版层次分明、结构清晰；书的末尾配有习题解答。

本书适合大学高年级学生以及研究生自学或作为教材使用。

Bryan P. Rynne, Martin A. Youngson
Linear Functional Analysis
EISBN: 1-85233-257-3

北京市版权局著作权合同登记号　图字：01-2004-3589

图书在版编目(CIP)数据

线性泛函分析 = Linear Functional Analysis / 瑞利（Rynne, B. P.），杨森（Youngson, M. A.）著. —影印本. —北京：清华大学出版社，2005.12
ISBN 7-302-12139-7

Ⅰ. 线…　Ⅱ. ①瑞…　②杨…　Ⅲ. 泛函分析－高等学校－教材－英文　Ⅳ. O177

中国版本图书馆 CIP 数据核字（2005）第 136637 号

出 版 者：清华大学出版社　　地　　址：北京清华大学学研大厦
http://www.tup.com.cn　　邮　　编：100084
社 总 机：(010) 6277 0175　　客户服务：(010) 6277 6969
责任编辑：陈朝晖
印 装 者：北京鑫海金澳胶印有限公司
发 行 者：新华书店总店北京发行所
开　　本：175×245　印张：17.75
版　　次：2005 年 12 月第 1 版　2005 年 12 月第 1 次印刷
书　　号：ISBN 7-302-12139-7/O・506
印　　数：1~3000
定　　价：29.00 元

Preface

This book provides an introduction to the ideas and methods of linear functional analysis at a level appropriate to the final year of an undergraduate course at a British university. The prerequisites for reading it are a standard undergraduate knowledge of linear algebra and real analysis (including the theory of metric spaces).

Part of the development of functional analysis can be traced to attempts to find a suitable framework in which to discuss differential and integral equations. Often, the appropriate setting turned out to be a vector space of real or complex-valued functions defined on some set. In general, such a vector space is infinite-dimensional. This leads to difficulties in that, although many of the elementary properties of finite-dimensional vector spaces hold in infinite-dimensional vector spaces, many others do not. For example, in general infinite-dimensional vector spaces there is no framework in which to make sense of analytic concepts such as convergence and continuity. Nevertheless, on the spaces of most interest to us there is often a *norm* (which extends the idea of the length of a vector to a somewhat more abstract setting). Since a norm on a vector space gives rise to a metric on the space, it is now possible to do analysis in the space. As real or complex-valued functions are often called *functionals*, the term *functional analysis* came to be used for this topic.

We now briefly outline the contents of the book. In Chapter 1 we present (for reference and to establish our notation) various basic ideas that will be required throughout the book. Specifically, we discuss the results from elementary linear algebra and the basic theory of metric spaces which will be required in later chapters. We also give a brief summary of the elements of the theory of Lebesgue measure and integration. Of the three topics discussed in this introductory chapter, Lebesgue integration is undoubtedly the most technically difficult and the one which the prospective reader is least likely to have encountered before. Unfortunately, many of the most important spaces which arise in functional analysis are spaces of integrable functions, and it is necessary to use the Lebesgue integral to overcome various drawbacks of the elementary Riemann integral, commonly taught in real analysis courses. The reader who has not met Lebesgue integration before can still read this book by accepting that an integration process exists which coincides with the Riemann integral when this is defined, but extends to a larger class of functions, and which has the properties described in Section 1.3.

In Chapter 2 we discuss the fundamental concept of functional analysis, the *normed vector space*. As mentioned above, a norm on a vector space is simply an extension of the idea of the length of a vector to a rather more abstract setting. Via an associated metric, the norm is behind all the discussion of convergence and continuity in vector spaces in this book. The basic properties of normed vector spaces are described in this chapter. In particular we begin the study of *Banach spaces* which are complete normed vector spaces.

In finite dimensions, in addition to the length of a vector, the angle between two vectors is also used. To extend this to more abstract spaces the idea of an *inner product* on a vector space is introduced. This generalizes the well-known "dot product" used in $\mathbb{R}^3$. *Inner product spaces*, which are vector spaces possessing an inner product, are discussed in Chapter 3. Every inner product space is a normed space and, as in Chapter 2, we find that the most important inner product spaces are those which are complete. These are called *Hilbert spaces*.

Having discussed various properties of infinite-dimensional vector spaces the next step is to look at linear transformations between these spaces. The most important linear transformations are the continuous ones, and these will be called linear *operators*. In Chapter 4 we describe general properties of linear operators between normed vector spaces. Any linear transformation between finite-dimensional vector spaces is automatically continuous so questions relating to the continuity of the transformation can safely be ignored (and usually are). However, when the spaces are infinite-dimensional this is certainly not the case and the continuity, or otherwise, of individual linear transformations must be studied much more carefully. In addition, we investigate the algebraic properties of the entire set of all linear operators between given normed vector spaces. Finally, for some linear operators it is possible to define an inverse operator, and we conclude the chapter with a characterization of the invertibility of an operator.

In Chapter 5 we specialize the discussion of linear operators to those acting between Hilbert spaces. The additional structure of these spaces means that we can define the *adjoint* of a linear operator and hence the particular classes of *self-adjoint* and *unitary* operators which have especially nice properties. We also introduce the *spectrum* of linear operators acting on a Hilbert space. The spectrum of a linear operator is a generalization of the set of eigenvalues of a matrix, which is a well-known concept in finite-dimensional linear algebra.

As we have already remarked, there are many significant differences between the theory of linear transformations in finite and infinite dimensions. However, for the class of *compact* operators a great deal of the theory carries over from finite to infinite dimensions. The properties of these particular operators are discussed in detail in Chapter 6. In particular, we study compact, self-adjoint operators on Hilbert spaces, and their spectral properties.

Finally, in Chapter 7, we use the results of the preceding chapters to discuss two extremely important areas of application of functional analysis, namely integral and differential equations. As we remarked above, the study of these equations was one of the main early influences and driving forces in the growth and development of functional analysis, so it forms a fitting conclusion to this book. Nowadays, functional analysis has applications to a vast range of areas of mathematics, but limitations of space preclude us from studying further applications.

A large number of exercises are included, together with complete solutions. Many of these exercises are relatively simple, while some are considerably less so. It is strongly recommended that the student should at least attempt most of these questions before looking at the solution. This is the only way to really learn any branch of mathematics.

There is a World Wide Web site associated with this book, at the URL

`http://www.ma.hw.ac.uk/~bryan/lfa_book.html`

This site contains links to sites on the Web which give some historical background to the subject, and also contains a list of any significant misprints which have been found in the book.

Contents

1. Preliminaries 1
- 1.1 Linear Algebra 2
- 1.2 Metric Spaces 11
- 1.3 Lebesgue Integration 20

2. Normed Spaces 31
- 2.1 Examples of Normed Spaces 31
- 2.2 Finite-dimensional Normed Spaces 39
- 2.3 Banach Spaces 45

3. Inner Product Spaces, Hilbert Spaces 51
- 3.1 Inner Products 51
- 3.2 Orthogonality 60
- 3.3 Orthogonal Complements 65
- 3.4 Orthonormal Bases in Infinite Dimensions 72
- 3.5 Fourier Series 82

4. Linear Operators 87
- 4.1 Continuous Linear Transformations 87
- 4.2 The Norm of a Bounded Linear Operator 96
- 4.3 The Space $B(X, Y)$ and Dual Spaces 104
- 4.4 Inverses of Operators 111

5. Linear Operators on Hilbert Spaces 123
- 5.1 The Adjoint of an Operator 123
- 5.2 Normal, Self-adjoint and Unitary Operators 132
- 5.3 The Spectrum of an Operator 139

5.4 Positive Operators and Projections 148

6. Compact Operators 161
6.1 Compact Operators 161
6.2 Spectral Theory of Compact Operators 172
6.3 Self-adjoint Compact Operators 182

7. Integral and Differential Equations 191
7.1 Fredholm Integral Equations 191
7.2 Volterra Integral Equations 201
7.3 Differential Equations 203
7.4 Eigenvalue Problems and Green's Functions 208

8. Solutions to Exercises 221

Further Reading 265

References 267

Notation Index 269

Index 271

1 Preliminaries

To a certain extent, functional analysis can be described as infinite-dimensional linear algebra combined with analysis, in order to make sense of ideas such as convergence and continuity. It follows that we will make extensive use of these topics, so in this chapter we briefly recall and summarize the various ideas and results which are fundamental to the study of functional analysis. We must stress, however, that this chapter only attempts to review the material and establish the notation that we will use. We do not attempt to motivate or explain this material, and any reader who has not met this material before should consult an appropriate textbook for more information.

Section 1.1 discusses the basic results from linear algebra that will be required. The material here is quite standard although, in general, we do not make any assumptions about finite-dimensionality except where absolutely necessary. Section 1.2 discusses the basic ideas of metric spaces. Metric spaces are the appropriate setting in which to discuss basic analytical concepts such as convergence of sequences and continuity of functions. The ideas are a natural extension of the usual concepts encountered in elementary courses in real analysis. In general metric spaces no other structure is imposed beyond a metric, which is used to discuss convergence and continuity. However, the essence of functional analysis is to consider vector spaces (usually infinite-dimensional) which are metric spaces and to study the interplay between the algebraic and metric structures of the spaces, especially when the spaces are complete metric spaces.

An important technical tool in the theory is Lebesgue integration. This is because many important vector spaces consist of sets of integrable functions.

In order for desirable metric space properties, such as completeness, to hold in these spaces it is necessary to use Lebesgue integration rather than the simpler Riemann integration usually discussed in elementary analysis courses. Of the three topics discussed in this introductory chapter, Lebesgue integration is undoubtedly the most technically difficult and the one which the prospective student is most likely to have not encountered before. In this book we will avoid arcane details of Lebesgue integration theory. The basic results which will be needed are described in Section 1.3, without any proofs. For the reader who is unfamiliar with Lebesgue integration and who does not wish to embark on a prolonged study of the theory, it will be sufficient to accept that an integration process exists which applies to a broad class of "Lebesgue integrable" functions and has the properties described in Section 1.3, most of which are obvious extensions of corresponding properties of the Riemann integral.

1.1 Linear Algebra

Throughout the book we have attempted to use standard mathematical notation wherever possible. Basic to the discussion is standard set theoretic notation and terminology. Details are given in, for example, [7]. Sets will usually be denoted by upper case letters, $X, Y, \ldots$, while elements of sets will be denoted by lower case letters, $x, y, \ldots$. The usual set theoretic operations will be used: $\in$, $\subset$, $\cup$, $\cap$, $\emptyset$ (the empty set), $\times$ (Cartesian product), $X \setminus Y = \{x \in X : x \notin Y\}$.

The following standard sets will be used,

$$\mathbb{R} = \text{the set of real numbers},$$
$$\mathbb{C} = \text{the set of complex numbers},$$
$$\mathbb{N} = \text{the set of positive integers } \{1, 2, \ldots\}.$$

The sets $\mathbb{R}$ and $\mathbb{C}$ are algebraic *fields*. These fields will occur throughout the discussion, associated with vector spaces. Sometimes it will be crucial to be specific about which of these fields we are using, but when the discussion applies equally well to both we will simply use the notation $\mathbb{F}$ to denote either set. The real and imaginary parts of a complex number z will be denoted by $\Re\mathrm{e}\, z$ and $\Im\mathrm{m}\, z$ respectively, while the complex conjugate will be denoted $\overline{z}$.

For any $k \in \mathbb{N}$ we let $\mathbb{F}^k = \mathbb{F} \times \ldots \times \mathbb{F}$ (the Cartesian product of k copies of $\mathbb{F}$). Elements of $\mathbb{F}^k$ will written in the form $x = (x_1, \ldots, x_k)$, $x_j \in \mathbb{F}$, $j = 1, \ldots, k$.

For any two sets X and Y, the notation $f : X \to Y$ will denote a function or mapping from X into Y. The set X is the *domain* of f and Y is the *codomain*. If $A \subset X$ and $B \subset Y$, we use the notation

$$f(A) = \{f(x) : x \in A\}, \quad f^{-1}(B) = \{x \in X : f(x) \in B\}.$$

If Z is a third set and $g : Y \to Z$ is another function, we define the *composition* of g and f, written $g \circ f : X \to Z$, by

$$(g \circ f)(x) = g(f(x)),$$

for all $x \in X$.

We now discuss the essential concepts from linear algebra that will be required in later chapters. Most of this section should be familiar, at least in the finite-dimensional setting, see for example [1] or [5], or any other book on linear algebra. However, we do not assume here that any spaces are finite-dimensional unless explicitly stated.

Definition 1.1

A *vector space over* $\mathbb{F}$ is a non-empty set V together with two functions, one from $V \times V$ to V and the other from $\mathbb{F} \times V$ to V, denoted by $x + y$ and αx respectively, for all $x, y \in V$ and $\alpha \in \mathbb{F}$, such that, for any α, $\beta \in \mathbb{F}$ and any x, y, $z \in V$,

(a) $x + y = y + x, \quad x + (y + z) = (x + y) + z;$

(b) there exists a unique $0 \in V$ (independent of x) such that $x + 0 = x$;

(c) there exists a unique $-x \in V$ such that $x + (-x) = 0$;

(d) $1x = x, \quad \alpha(\beta x) = (\alpha\beta)x;$

(e) $\alpha(x + y) = \alpha x + \alpha y, \quad (\alpha + \beta)x = \alpha x + \beta x.$

If $\mathbb{F} = \mathbb{R}$ (respectively, $\mathbb{F} = \mathbb{C}$) then V is a *real* (respectively, *complex*) vector space. Elements of $\mathbb{F}$ are called *scalars*, while elements of V are called *vectors*. The operation $x + y$ is called *vector addition*, while the operation αx is called *scalar multiplication*.

Many results about vector spaces apply equally well to both real or complex vector spaces so if the type of a space is not stated explicitly then the space may be of either type, and we will simply use the term "vector space".

If V is a vector space with $x \in V$ and A, $B \subset V$, we use the notation,

$$x + A = \{x + a : a \in A\},$$

$$A + B = \{a + b : a \in A,\ b \in B\}.$$

Definition 1.2

Let V be a vector space. A non-empty set $U \subset V$ is a *linear subspace* of V if U is itself a vector space (with the same vector addition and scalar multiplication as in V). This is equivalent to the condition that

$$\alpha x + \beta y \in U, \quad \text{for all } \alpha, \beta \in \mathbb{F} \text{ and } x, y \in U$$

(which is called the *subspace test*).

Note that, by definition, vector spaces and linear subspaces are always non-empty, while general subsets of vector spaces which are not subspaces may be empty. In particular, it is a consequence of the vector space definitions that $0x = \mathbf{0}$, for all $x \in V$ (here, 0 is the scalar zero and $\mathbf{0}$ is the vector zero; except where it is important to distinguish between the two, both will be denoted by 0). Hence, any linear subspace $U \subset V$ must contain at least the vector $\mathbf{0}$, and the set $\{\mathbf{0}\} \subset V$ is a linear subspace.

Definition 1.3

Let V be a vector space, let $\mathbf{v} = \{v_1, \ldots, v_k\} \subset V$, $k \geq 1$, be a finite set and let $A \subset V$ be an arbitrary non-empty set.

(a) A *linear combination* of the elements of $\mathbf{v}$ is any vector of the form

$$x = \alpha_1 v_1 + \cdots + \alpha_k v_k \in V, \tag{1.1}$$

for any set of scalars $\alpha_1, \ldots, \alpha_k$.

(b) $\mathbf{v}$ is *linearly independent* if the following implication holds:

$$\alpha_1 v_1 + \ldots + \alpha_k v_k = 0 \quad \Rightarrow \quad \alpha_1 = \ldots = \alpha_k = 0.$$

(c) A is *linearly independent* if every finite subset of A is linearly independent. If A is not linearly independent then it is *linearly dependent.*

(d) The *span* of A (denoted $\operatorname{Sp} A$) is the set of all linear combinations of all finite subsets of A. This set is a linear subspace of V. Equivalently, $\operatorname{Sp} A$ is the intersection of the set of all linear subspaces of V which contain A. Thus, $\operatorname{Sp} A$ is the smallest linear subspace of V containing A (in the sense that if $A \subset B \subset V$ and B is a linear subspace of V then $\operatorname{Sp} A \subset B$).

(e) If $\mathbf{v}$ is linearly independent and $\operatorname{Sp} \mathbf{v} = V$, then $\mathbf{v}$ is called a *basis* for V. It can be shown that if V has such a (finite) basis then all bases of V have the same number of elements. If this number is k then V is said to be *k-dimensional* (or, more generally, *finite-dimensional*), and we write $\dim V = k$. If V does not have such a finite basis it is said to be *infinite-dimensional.*

(f) If $\mathbf{v}$ is a basis for V then any $x \in V$ can be written as a linear combination of the form (1.1), with a unique set of scalars α_j, $j = 1, \ldots, k$. These scalars (which clearly depend on x) are called the *components* of x with respect to the basis $\mathbf{v}$.

(g) The set $\mathbb{F}^k$ is a vector space over $\mathbb{F}$ and the set of vectors

$$\widehat{e}_1 = (1, 0, 0, \ldots, 0),\ \widehat{e}_2 = (0, 1, 0, \ldots, 0), \ldots,\ \widehat{e}_k = (0, 0, 0, \ldots, 1),$$

is a basis for $\mathbb{F}^k$. This notation will be used throughout the book, and this basis will be called the *standard basis* for $\mathbb{F}^k$.

We will sometimes write $\dim V = \infty$ when V is infinite-dimensional. However, this is simply a notational convenience, and should not be interpreted in the sense of ordinal or cardinal numbers (see [7]). In a sense, infinite-dimensional spaces can vary greatly in their "size"; see Section 3.4 for some further discussion of this.

Definition 1.4

Let V, W be vector spaces over $\mathbb{F}$. The Cartesian product $V \times W$ is a vector space with the following vector space operations. For any $\alpha \in \mathbb{F}$ and any $(x_j, y_j) \in V \times W$, $j = 1, 2$, let

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + x_2), \quad \alpha(x_1, y_1) = (\alpha x_1, \alpha y_1)$$

(using the corresponding vector space operations in V and W).

We next describe a typical construction of vector spaces consisting of functions defined on some underlying set.

Definition 1.5

Let S be a set and let V be a vector space over $\mathbb{F}$. We denote the set of functions $f : S \to V$ by $F(S, V)$. For any $\alpha \in \mathbb{F}$ and any $f, g \in F(S, V)$, we define functions $f + g$ and αf in $F(S, V)$ by

$$(f + g)(x) = f(x) + g(x), \quad (\alpha f)(x) = \alpha f(x),$$

for all $x \in S$ (using the vector space operations in V). With these definitions the set $F(S, V)$ is a vector space over $\mathbb{F}$.

Many of the vector spaces used in functional analysis are of the above form. From now on, whenever functions are added or multiplied by a scalar the process will be as in Definition 1.5. We note that the zero element in $F(S, V)$ is the function which is identically equal to the zero element of V. Also, if S contains infinitely many elements and $V \neq \{0\}$ then $F(S, V)$ is infinite-dimensional.

Example 1.6

If S is the set of integers $\{1, \ldots, k\}$ then the set $F(S, \mathbb{F})$ can be identified with the space $\mathbb{F}^k$ (by identifying an element $x \in \mathbb{F}^k$ with the function $f \in F(S, \mathbb{F})$ defined by $f(j) = x_j$, $1 \leq j \leq k$).

Often, in the construction in Definition 1.5, the set S is a vector space and only a subset of the set of all functions $f : S \to V$ is considered. In particular, in this case the most important functions to consider are those which preserve the linear structure of the vector spaces in the sense of the following definition.

Definition 1.7

Let V, W be vector spaces over the same scalar field $\mathbb{F}$. A function $T : V \to W$ is called a *linear transformation* (or *mapping*) if, for all α, $\beta \in \mathbb{F}$ and x, $y \in V$,

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y).$$

The set of all linear transformations $T : V \to W$ will be denoted by $L(V, W)$. With the scalar multiplication and vector addition defined in Definition 1.5 the set $L(V, W)$ is a vector space (it is a subspace of $F(V, W)$). When $V = W$ we abbreviate $L(V, V)$ to $L(V)$.

A particularly simple linear transformation in $L(V)$ is defined by $I_V(x) = x$, for $x \in V$. This is called the *identity* transformation on V (usually we use the notation I if it is clear what space the transformation is acting on).

Whenever we discuss linear transformations $T : V \to W$ it will be taken for granted, without being explicitly stated, that V and W are vector spaces over the same scalar field.

Since linear transformations are functions they can be composed (when they act on appropriate spaces). The following lemmas are immediate consequences of the definition of a linear transformation.

Lemma 1.8

Let V, W, X be vector spaces and $T \in L(V, W)$, $S \in L(W, X)$. Then the composition $S \circ T \in L(V, X)$.

Lemma 1.9

Let V be a vector space, R, S, $T \in L(V)$, and $\alpha \in \mathbb{F}$. Then:

(a) $R\circ(S\circ T) = (R\circ S)\circ T$;

(b) $R\circ(S+T) = R\circ S + R\circ T$;

(c) $(S+T)\circ R = S\circ R + T\circ R$;

(d) $I_V\circ T = T\circ I_V = T$;

(e) $(\alpha S)\circ T = \alpha(S\circ T) = S\circ(\alpha T)$.

These properties also hold for linear transformations between different spaces when the relevant operations make sense (for instance, (a) holds when $T \in L(V,W)$, $S \in L(W,X)$ and $R \in L(X,Y)$, for vector spaces V, W, X, Y).

The five properties listed in Lemma 1.9 are exactly the extra axioms which a vector space must satisfy in order to be an *algebra*. Since this is the only example of an algebra which we will meet in this book we will not discuss this further, but we note that an algebra is both a vector space and a *ring*, see [5].

When dealing with the composition of linear transformations S, T it is conventional to omit the symbol $\circ$ and simply write ST. Eventually we will do this, but for now we retain the symbol $\circ$.

The following lemma gives some further elementary properties of linear transformations.

Lemma 1.10

Let V, W be vector spaces and $T \in L(V,W)$.

(a) $T(0) = 0$.

(b) If U is a linear subspace of V then the set $T(U)$ is a linear subspace of W and $\dim T(U) \le \dim U$ (as either finite numbers or ∞).

(c) If U is a linear subspace of W then the set $\{x \in V : T(x) \in U\}$ is a linear subspace of V.

We can now state some standard terminology.

Definition 1.11

Let V, W be vector spaces and $T \in L(V,W)$.

(a) The *image* of T (often known as the *range* of T) is the subspace $\operatorname{Im} T = T(V)$; the *rank* of T is the number $r(T) = \dim(\operatorname{Im} T)$.

(b) The *kernel* of T (often known as the *null-space* of T) is the subspace $\operatorname{Ker} T = \{x \in V : T(x) = 0\}$; the *nullity* of T is the number $n(T) = \dim(\operatorname{Ker} T)$.

The rank and nullity, $r(T)$, $n(T)$, may have the value ∞.

(c) T has *finite rank* if $r(T)$ is finite.

(d) T is *one-to-one* if, for any $y \in W$, the equation $T(x) = y$ has at most one solution x.

(e) T is *onto* if, for any $y \in W$, the equation $T(x) = y$ has at least one solution x.

(f) T is *bijective* if, for any $y \in W$, the equation $T(x) = y$ has exactly one solution x (that is, T is both one-to-one and onto).

Lemma 1.12

Let V, W be vector spaces and $T \in L(V, W)$.

(a) T is one-to-one if and only if the equation $T(x) = 0$ has only the solution $x = 0$. This is equivalent to $\operatorname{Ker} T = \{0\}$ or $n(T) = 0$.

(b) T is onto if and only if $\operatorname{Im} T = W$. If $\dim W$ is finite this is equivalent to $r(T) = \dim W$.

(c) $T \in L(V, W)$ is bijective if and only if there exists a transformation $S \in L(W, V)$ which is bijective and $S \circ T = I_V$ and $T \circ S = I_W$.

If V is k-dimensional then

$$n(T) + r(T) = k$$

(in particular, $r(T)$ is necessarily finite, irrespective of whether W is finite-dimensional). Hence, if W is also k-dimensional then T is bijective if and only if $n(T) = 0$.

Related to the bijectivity, or otherwise, of a transformation T from a space to itself we have the following definition, which will be extremely important later.

Definition 1.13

Let V be a vector space and $T \in L(V)$. A scalar $\lambda \in \mathbb{F}$ is an *eigenvalue* of T if the equation $T(x) = \lambda x$ has a non-zero solution $x \in V$, and any such non-zero solution is an *eigenvector*. The subspace $\operatorname{Ker}(T - \lambda I) \subset V$ is called the *eigenspace* (corresponding to λ) and the *multiplicity* of λ is the number $m_\lambda = n(T - \lambda I)$.

Lemma 1.14

Let V be a vector space and let $T \in L(V)$. Let $\{\lambda_1, \ldots, \lambda_k\}$ be a set of distinct eigenvalues of T, and for each $1 \le j \le k$ let x_j be an eigenvector corresponding to λ_j. Then the set $\{x_1, \ldots, x_k\}$ is linearly independent.

Linear transformations between finite-dimensional vector spaces are closely related to matrices. For any integers $m, n \ge 1$, let $M_{mn}(\mathbb{F})$ denote the set of all $m \times n$ matrices with entries in $\mathbb{F}$. A typical element of $M_{mn}(\mathbb{F})$ will be written as $[a_{ij}]$ (or $[a_{ij}]_{mn}$ if it is necessary to emphasize the size of the matrix). Any matrix $C = [c_{ij}] \in M_{mn}(\mathbb{F})$ induces a linear transformation $T_C \in L(\mathbb{F}^n, \mathbb{F}^m)$ as follows: for any $x \in \mathbb{F}^n$, let $T_C x = y$, where $y \in \mathbb{F}^m$ is defined by

$$y_i = \sum_{j=1}^{n} c_{ij} x_j, \quad 1 \le i \le m.$$

Note that, if we were to regard x and y as column vectors then this transformation corresponds to standard matrix multiplication. However, mainly for notational purposes, it is generally convenient to regard elements of $\mathbb{F}^k$ as row vectors. This convention will always be used below, except when we specifically wish to perform computations of matrices acting on vectors, and then it will be convenient to use column vector notation.

On the other hand, if U and V are finite-dimensional vector spaces then a linear transformation $T \in L(U, V)$ can be represented in terms of a matrix. To fix our notation we briefly review this representation (see Chapter 7 of [1] for further details). Suppose that U is n-dimensional and V is m-dimensional, with bases $\mathbf{u} = \{u_1, \ldots, u_n\}$ and $\mathbf{v} = \{v_1, \ldots, v_m\}$ respectively. Any vector $a \in U$ can be represented in the form

$$a = \sum_{j=1}^{n} \alpha_j u_j,$$

for a unique collection of scalars $\alpha_1, \ldots, \alpha_n$. We define the column matrix

$$A = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in M_{n1}(\mathbb{F}).$$

The mapping $a \to A$ is a bijective linear transformation from U to $M_{n1}(\mathbb{F})$, that is, there is a one-to-one correspondence between vectors $a \in U$ and column matrices $A \in M_{n1}(\mathbb{F})$. There is a similar correspondence between vectors $b \in V$ and column matrices $B \in M_{m1}(\mathbb{F})$. Now, for any $1 \le j \le n$, the vector Tu_j has the representation

$$Tu_j = \sum_{i=1}^{m} \tau_{ij} v_i,$$