CHAPTER 1

Fundamentals

This first chapter aims to introduce the notion of an abstract linear space to those who think of vectors as arrays of components. I want to point out that the class of abstract linear spaces is no larger than the class of spaces whose elements are arrays. So what is gained by this abstraction?

First of all, the freedom to use a single symbol for an array; this way we can think of vectors as basic building blocks, unencumbered by components. The abstract view leads to simple, transparent proofs of results.

More to the point, the elements of many interesting vector spaces are not presented in terms of components. For instance, take a linear ordinary differential equation of degree n; the set of its solutions form a vector space of dimension n, yet they are not presented as arrays.

Even if the elements of a vector space are presented as arrays of numbers, the elements of a subspace of it may not have a natural description as arrays. Take, for instance, the subspace of all vectors whose components add up to zero.

Last but not least, the abstract view of vector spaces is indispensable for infinite-dimensional spaces; even though this text is strictly about finite-dimensional spaces, it is a good preparation for functional analysis.

Linear algebra abstracts the two basic operations with vectors: the addition of vectors, and their multiplication by numbers (scalars). It is astonishing that on such slender foundations an elaborate structure can be built, with romanesque, gothic, and baroque aspects. It is even more astounding that linear algebra has not only the right theorems but also the right language for many mathematical topics, including applications of mathematics.

A *linear space* X over a *field* K is a mathematical object in which two operations are defined:

Addition, denoted by $+$, as in

$$x + y \tag{1}$$

Linear Algebra and Its Applications, Second Edition, by Peter D. Lax

and assumed to be *commutative*:

$$x + y = y + x, \tag{2}$$

and *associative*:

$$x + (y + z) = (x + y) + z, \tag{3}$$

and to form a *group*, with the neutral element denoted as 0:

$$x + 0 = x. \tag{4}$$

The inverse of addition is denoted by $-$:

$$x + (-x) \equiv x - x = 0. \tag{5}$$

EXERCISE 1. Show that the zero of vector addition is unique.

The second operation is *multiplication* of elements of X *by elements* k *of the field* K:

$$kx.$$

The result of this multiplication is a vector, that is, an element of X.

Multiplication by elements of K is assumed to be *associative*:

$$k(ax) = (ka)x \tag{6}$$

and *distributive*:

$$k(x + y) = kx + ky, \tag{7}$$

as well as

$$(a + b)x = ax + bx. \tag{8}$$

We assume that multiplication by the *unit* of K, denoted as 1, acts as the identity:

$$1x = x. \tag{9}$$

These are the *axioms* of linear algebra. We proceed to draw some deductions:

Set $b = 0$ in (8); it follows from Exercise 1 that for all x

$$0x = 0. \tag{10}$$

Set $a = 1, b = -1$ in (8); using (9) and (10) we deduce that for all x

$$(-1)x = -x.$$

EXERCISE 2. Show that the vector with all components zero serves as the zero element of classical vector addition.

In this analytically oriented text the field K will be either the field $\mathbb{R}$ of real numbers or the field $\mathbb{C}$ of complex numbers.

An interesting example of a linear space is the set of all functions $x(t)$ that satisfy the differential equation

$$\frac{d^2}{dt^2}x + x = 0.$$

The sum of two solutions is again a solution, and so is the constant multiple of one. This shows that the set of solutions of this differential equation form a linear space.

Solutions of this equation describe the motion of a mass connected to a fixed point by a spring. Once the initial position $x(0) = p$ and initial velocity $\frac{d}{dt}x(0) = v$ are given, the motion is completely determined for all t. So solutions can be described by a pair of numbers (p, v).

The relation between the two descriptions is *linear*; that is, if (p, v) are the initial data of a solution $x(t)$, and (q, w) the initial data of another solution $y(t)$, then the initial data of the solution $x(t) + y(t)$ are $(p + q, v + w) = (p, v) + (q, w)$. Similarly, the initial data of the solution $kx(t)$ are $(kp, kv) = k(p, v)$.

This kind of relation has been abstracted into the notion of *isomorphism.*

Definition. A one-to-one correspondence between two linear spaces over the same field that maps sums into sums and scalar multiples into scalar multiples is called an *isomorphism*.

Isomorphism is a basic notion in linear algebra. Isomorphic linear spaces are indistinguishable by means of operations available in linear spaces. Two linear spaces that are presented in very different ways can be, as we have seen, isomorphic.

Examples of Linear Spaces. **(i)** Set of all row vectors: $(a_1, \ldots, a_n), a_j$ in K; addition, multiplication defined componentwise. This space is denoted as K^n.

(ii) Set of all real-valued functions $f(x)$ defined on the real line, $K = \mathbb{R}$.

(iii) Set of all functions with values in K, defined on an arbitrary set S.

(iv) Set of all polynomials of degree less than n with coefficients in K.

EXERCISE 3. Show that (i) and (iv) are isomorphic.

EXERCISE 4. Show that if S has n elements, (i) and (iii) are isomorphic.

EXERCISE 5. Show that when $K = \mathbb{R}$, (iv) is isomorphic with (iii) when S consists of n distinct points of $\mathbb{R}$.

Definition. A subset Y of a linear space X is called a *subspace* if sums and scalar multiples of elements of Y belong to Y.

Examples of Subspaces. (**a**) X as in Example (i), Y the set of vectors $(0, a_2, \dots, a_{n-1}, 0)$ whose first and last component is zero.

(**b**) X as in Example (ii), Y the set of all periodic functions with period π.

(**c**) X as in Example (iii), Y the set of constant functions on S.

(**d**) X as in Example (iv), Y the set of all even polynomials.

Definition. The *sum* of two subsets Y and Z of a linear space X, denoted as $Y + Z$, is the set of all vectors of form $y + z$, y in Y, z in Z.

EXERCISE 6. Prove that $Y + Z$ is a linear subspace of X if Y and Z are.

Definition. The *intersection* of two subsets Y and Z of a linear space X, denoted as $Y \cap Z$, consists of all vectors x that belong to both Y and Z.

EXERCISE 7. Prove that if Y and Z are linear subspaces of X, so is $Y \cap Z$.

EXERCISE 8. Show that the set $\{0\}$ consisting of the zero element of a linear space X is a subspace of X. It is called the *trivial subspace.*

Definition. *A linear combination of* j vectors $x_1, \dots, x_j$ of a linear space is a vector of the form

$$k_1 x_1 + \cdots + k_j x_j, \quad k_1, \dots, k_j \in K.$$

EXERCISE 9. Show that the set of *all* linear combinations of $x_1, \dots, x_j$ is a subspace of X, and that it is the smallest subspace of X containing $x_1, \dots, x_j$. This is called the *subspace spanned* by $x_1, \dots, x_j$.

Definition. A set of vectors $x_1, \dots, x_m$ in X *span* the whole space X if every x in X can be expressed as a linear combination of $x_1, \dots, x_m$.

Definition. The vectors $x_1, \dots, x_j$ are called *linearly dependent* if there is a nontrivial linear relation between them, that is, a relation of the form

$$k_1 x_1 + \cdots + k_j x_j = 0,$$

where not all $k_1, \dots, k_j$ are zero.

Definition. A set of vectors $x_1, \ldots, x_j$ that are not linearly dependent is called *linearly independent.*

EXERCISE 10. Show that if the vectors $x_1, \ldots, x_j$ are linearly independent, then none of the x_i is the zero vector.

Lemma 1. Suppose that the vectors $x_1, \ldots, x_n$ span a linear space X and that the vectors $y_1, \ldots, y_j$ in X are linearly independent. Then

$$j \leq n.$$

Proof. Since $x_1, \ldots, x_n$ span X, every vector in X can be written as a linear combination of $x_1, \ldots, x_n$. In particular, y_1:

$$y_1 = k_1 x_1 + \cdots + k_n x_n.$$

Since $y_1 \neq 0$ (see Exercise 10), not all k are equal to 0, say $k_i \neq 0$. Then x_i can be expressed as a linear combination of y_1 and the remaining x_s. So the set consisting of the x's, with x_i replaced by y_1 span X. If $j \geq n$, repeat this step $n - 1$ more times and conclude that $y_1, \ldots, y_n$ span X: if $j > n$, this contradicts the linear independence of the y's for then y_{n+1} is a linear combination of $y_1, \ldots, y_n$. $\square$

Definition. A finite set of vectors which span X and are linearly independent is called a *basis* for X.

Lemma 2. A linear space X which is spanned by a finite set of vectors $x_1, \ldots, x_n$ has a basis.

Proof. If $x_1, \ldots, x_n$ are linearly dependent, there is a nontrivial relation between them; from this one of the x_i can be expressed as a linear combination of the rest. So we can drop that x_i. Repeat this step until the remaining x_j are linear independent: they still span X, and so they form a basis. $\square$

Definition. A linear space X is called *finite dimensional* if it has a basis.

A finite-dimensional space has many, many bases. When the elements of the space are represented as arrays with n components, we give preference to the special basis consisting of the vectors that have one component equal to 1, while all the others equal 0.

Theorem 3. All bases for a finite-dimensional linear space X contain the same number of vectors. This number is called the dimension of X and is denoted as

$$\dim X.$$

Proof. Let $x_1, \ldots, x_n$ be one basis, and let $y_1, \ldots, y_m$ be another. By Lemma 1 and the definition of basis we conclude that $m \leq n$, and also $n \leq m$. So we conclude that n and m are equal. □

We define the dimension of the trivial space consisting of the single element 0 to be zero.

Theorem 4. Every linearly independent set of vectors $y_1, \ldots, y_j$ in a finite-dimensional linear space X can be completed to a basis of X.

Proof. If $y_1, \ldots, y_j$ do not span X, there is some x_1 that cannot be expressed as a linear combination of $y_1, \ldots, y_j$. Adjoin this x_1 to the y's. Repeat this step until the y's span X. This will happen in less than n steps, $n = \dim X$, because otherwise X would contain more than n linearly independent vectors, impossible for a space of dimension n. □

Theorem 4 illustrates the many different ways of forming a basis for a linear space.

Theorem 5. (**a**) Every subspace Y of a finite-dimensional linear space X is finite dimensional.

(**b**) Every subspace Y has a complement in X, that is, another subspace Z such that every vector x in X can be decomposed uniquely as

$$x = y + z, \qquad y \text{ in } Y, z \text{ in } Z. \tag{11}$$

Furthermore

$$\dim X = \dim Y + \dim Z. \tag{11$'$}$$

Proof. We can construct a basis in Y by starting with any nonzero vector y_1, and then adding another vector y_2 and another, as long as they are linearly independent. According to Lemma 1, there can be no more of these y_i than the dimension of X. A maximal set of linearly independent vectors $y_1, \ldots, y_j$ in Y spans Y, and so forms a basis of Y. According to Theorem 4, this set can be completed to form a basis of X by adjoining $Z_{j+1}, \ldots, Z_n$. Define Z as the space spanned by $Z_{j+1}, \ldots, Z_n$; clearly Y and Z are complements, and

$$\dim X = n = j + (n - j) = \dim Y + \dim Z. \qquad \square$$

Definition. X is said to be the *direct sum of* two subspaces Y and Z that are complements of each other. More generally X is said to be the direct sum of its subspaces $Y_1, \ldots, Y_m$ if every x in X can be expressed *uniquely* as

$$x = y_1 + \cdots + y_m, \qquad y_j \text{ in } Y_j, \tag{12}$$

This relation is denoted as

$$X = Y_1 \oplus \cdots \oplus Y_m.$$

EXERCISE 11. Prove that if X is finite dimensional and the direct sum of $Y_1, \ldots, Y_m$, then

$$\dim X = \sum \dim Y_j. \tag{12$'$}$$

Definition. An $(n-1)$-dimensional subspace of an n-dimensional space is called a hyperplane.

EXERCISE 12. Show that every finite-dimensional space X over K is isomorphic to $K^n, n = \dim X$. Show that this isomorphism is not unique when n is >1.

Since every n-dimensional linear space over K is isomorphic to K^n, it follows that *two linear spaces over the same field and of the same dimension are isomorphic.*

Note: There are many ways of forming such an isomorphism; it is not unique.

The concept of congruence modulo a subspace, defined below, is a very useful tool.

Definition. For X a linear space, Y a subspace, we say that two vectors x_1, x_2 in X are *congruent modulo Y*, denoted

$$x_1 \equiv x_2 \bmod Y,$$

if $x_1 - x_2 \in Y$. Congruence mod Y is an equivalence relation, that is, it is

(i) symmetric: if $x_1 \equiv x_2$, then $x_2 \equiv x_1$.
(ii) reflexive: $x \equiv x$ for all x in X.
(iii) transitive: if $x_1 \equiv x_2, x_2 \equiv x_3$, then $x_1 \equiv x_3$.

EXERCISE 13. Prove (i)–(iii) above. Show furthermore that if $x_1 \equiv x_2$, then $kx_1 \equiv kx_2$ for every scalar k.

We can divide elements of X into *congruence classes* mod Y. The congruence class containing the vector x is the set of all vectors congruent with X; we denote it by $\{x\}$.

EXERCISE 14. Show that two congruence classes are either identical or disjoint.

The set of congruence classes can be made into a linear space by defining addition and multiplication by scalars, as follows:

$$\{x\} + \{z\} = \{x + z\}$$

and

$$k\{x\} = \{kx\}.$$

That is, the sum of the congruence class containing x and the congruence class containing z is the class containing $x + z$. Similarly for multiplication by scalars.

EXERCISE 15. Show that the above definition of addition and multiplication by scalars is independent of the choice of representatives in the congruence class.

The linear space of congruence classes defined above is called the *quotient space* of X mod Y and is denoted as

$$X(\text{mod } Y) \quad \text{or} \quad X/Y.$$

The following example is illuminating: Take X to be the linear space of all row vectors $(a_1, \ldots, a_n)$ with n components, and take Y to be all vectors $y = (0, 0, a_3, \ldots, a_n)$ whose first two components are zero. Then two vectors are congruent mod Y iff their first two components are equal. Each equivalence class can be represented by a vector with two components, the common components of all vectors in the equivalence class.

This shows that forming a quotient space amounts to throwing away information contained in those components that pertain to Y. This is a very useful simplification when we do not need the information contained in the neglected components.

The next result shows the usefulness of quotient spaces for counting the dimension of a subspace.

Theorem 6. Y is a subspace of a finite-dimensional linear space X; then

$$\dim Y + \dim(X/Y) = \dim X. \tag{13}$$

Proof. Let $y_1, \ldots, y_j$ be a basis for Y, $j = \dim Y$. According to Theorem 4, this set can be completed to form a basis for X by adjoining $x_{j+1}, \ldots, x_n, n = \dim X$. We claim that

$$\{x_{j+1}\}, \ldots, \{x_n\} \tag{13$'$}$$

form a basis for X/Y. To show this we have to verify two properties of the cosets (13)$'$:

(i) They span X/Y.

(ii) They are linearly independent.

(i) Since $y_1, \ldots, x_n$ form a basis for X, every x in X can be expressed as

$$x = \sum a_i y_i + \sum b_k x_k.$$

It follows that

$$\{x\} = \sum b_k \{x_k\}.$$

(ii) Suppose that

$$\sum c_k \{x_k\} = 0.$$

This means that

$$\sum c_k x_k = y, \qquad y \text{ in } Y.$$

Express y as $\sum d_i y_i$; we get

$$\sum c_k x_k - \sum d_i y_i = 0.$$

Since $y_1, \ldots, x_n$ form a basis, they are linearly independent, and so all the c_k and d_i are zero.

It follows that

$$\dim X/Y = \# \text{ of } x_k = n - j.$$

So

$$\dim Y + \dim X/Y = j + n - j = n = \dim X. \qquad \square$$

EXERCISE 16. Denote by X the linear space of all polynomials $p(t)$ of degree $< n$, and denote by Y the set of polynomials that are zero at $t_1, \ldots, t_j$, $j < n$.

(i) Show that Y is a subspace of X.
(ii) Determine dim Y.
(iii) Determine dim X/Y.

The following corollary is a consequence of Theorem 6.

Corollary 6′. A subspace Y of a finite-dimensional linear space X whose dimension is the same as the dimension of X is all of X.

EXERCISE 17. Prove Corollary $6'$.

Theorem 7. Suppose X is a finite-dimensional linear space, U and V two subspaces of X such that X is the sum of U and V:

$$X = U + V.$$

Denote by W the intersection of U and V:

$$W = U \cap V.$$

Then

$$\dim X = \dim U + \dim V - \dim W. \tag{14}$$

Proof. When the intersection W of U and V is the trivial space $\{0\}, \dim W = 0$, and (14) is relation $(11)'$ of Theorem 5. We show now how to use the notion of quotient space to reduce the general case to the simple case $\dim W = 0$.

Define $U_0 = U/W, V_0 = V/W$; then $U_0 \cap V_0 = \{0\}$, and so $X_0 = X/W$ satisfies

$$X_0 = U_0 + V_0.$$

So according to $(11)'$,

$$\dim X_0 = \dim U_0 + \dim V_0. \tag*{$(14)'$}$$

Applying (13) of Theorem 6 three times, we get

$$\dim X_0 = \dim X - \dim W, \qquad \dim U_0 = \dim U - \dim W,$$
$$\dim V_0 = \dim V - \dim W.$$

Setting this into relation $(14)'$ gives (14). $\square$

Definition. The *Cartesian sum* of two linear spaces over the same field is the set of pairs

$$(x_1, x_2); \quad x_1 \text{ in } X_1, x_2 \text{ in } X_2,$$

where addition and multiplication by scalars is defined componentwise. The direct sum is denoted as

$$X_1 \oplus X_2.$$

It is easy to verify that $X_1 \oplus X_2$ is indeed a linear space.

EXERCISE 18. Show that

$$\dim X_1 \oplus X_2 = \dim X_1 + \dim X_2.$$

EXERCISE 19. X a linear space, Y a subspace. Show that $Y \oplus X/Y$ is isomorphic to X.

Note: The most frequently occurring linear spaces in this text are our old friends $\mathbb{R}^n$ and $\mathbb{C}^n$, the spaces of vectors $(a_1, \ldots, a_n)$ with n real, respectively complex, components.

So far the only means we have for showing that a linear space X is finite dimensional is to find a finite set of vectors that span it. In Chapter 7 we present another, powerful criterion for a Euclidean space to be finite dimensional. In Chapter 14 we extend this criterion to all normed linear spaces.

We have been talking about sets of vectors being linearly dependent or independent, but have given no indication how to decide which is the case. Here is an example:

Decide if the four vectors

$$\begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 2 \\ 3 \end{pmatrix}$$

are linearly dependent or not. That is, are there four numbers k_1, k_2, k_3, k_4, not all zero, such that

$$k_1 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} + k_2 \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix} + k_3 \begin{pmatrix} 2 \\ 1 \\ 1 \\ 3 \end{pmatrix} + k_4 \begin{pmatrix} 2 \\ -1 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}?$$

This vector equation is equivalent to four scalar equations:

$$\begin{aligned} k_1 + k_2 + 2k_3 + 2k_4 &= 0, \\ k_1 - k_2 + k_3 - k_4 &= 0, \\ k_2 + k_3 &= 0, \\ k_1 + k_2 + 3k_3 + 3k_4 &= 0. \end{aligned} \tag{15}$$

The study of such systems of linear equations is the subject of Chapters 3 and 4. There we describe an algorithm for finding all solutions of such systems of equations.

EXERCISE 20. Which of the following sets of vectors $x = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$ are a subspace of $\mathbb{R}^n$? Explain your answer.

(**a**) All x such that $x_1 \geq 0$.
(**b**) All x such that $x_1 + x_2 = 0$.
(**c**) All x such that $x_1 + x_2 + 1 = 0$.
(**d**) All x such that $x_1 = 0$.
(**e**) All x such that x_1 is an integer.

EXERCISE 21. Let U, V, and W be subspaces of some finite-dimensional vector space X. Is the statement

$$\begin{aligned}\dim(U + V + W) = {} & \dim U + \dim V + \dim W - \dim(U \cap V) - \dim(U \cap W) \\ & - \dim(V \cap W) + \dim(U \cap V \cap W),\end{aligned}$$

true or false? If true, prove it. If false, provide a counterexample.