Foundations of Linear Algebra (preliminary)

David A. Pierce
pierced@member.ams.org
http://www.math.metu.edu.tr/~dpierce/

21 February 2001

Contents

0 Introduction

Linear algebra takes some of the concepts of high-school algebra and broadens their application, in such a way that new concepts arise. For example, linear algebra gives a way to define *dimension* and to exhibit spaces of all finite dimensions and of infinite dimensions. It has applications to subjects beyond algebra, such as the solution of differential equations.

These notes are intended as a terse summary of basic linear algebra and its background; as such they are but a supplement to a human lecturer or a textbook. They are fairly theoretical. Comments and corrections are welcome; the writing and editing is ongoing.

High-school algebra is largely a study of *real numbers* under the operations of *addition* and *multiplication*. Having studied high-school algebra, one should be able to apply the following facts and conventions almost instinctively. Real numbers themselves are generally symbolized by (plain-face, italic) letters, such as a, b, c or x, y, z. To an ordered pair (x, y) of real numbers x and y, the operation of addition assigns a *sum*, usually symbolized $x + y$. So, addition is a *binary* or *2-ary* operation. Multiplication assigns to (x, y) the *product* symbolized by $x \cdot y$ or xy; thus multiplication is also a binary operation.

There are particular real numbers 0 and 1, and on real numbers there are additional operations, derived from addition and multiplication. For every real number x there is an *additive inverse* or *negative*, $-x$. The operation of additive inversion is thus a *unary* or *1-ary* operation. In particular, there is a real number -1, and $-x$ is the product of -1 and x. There is a binary operation of *subtraction*, assigning to (x, y) the *difference* $x - y$, which is the sum of $-y$ and x.

For every real number x different from 0, there is a *multiplicative inverse* or *reciprocal*, x^{-1}; its product with x is 1. There is a *partial* binary operation of *division*, assigning to the pair (x, y) the *quotient* x/y, provided $y \neq 0$. This quotient is the product of y^{-1} and x, and is written also

$$\frac{x}{y}.$$

Note that $x^{-1} = 1/x$.

When operations are combined in an expression, one needs to know which to apply first. A standard (although not unique) convention is to find results in this order:

1. reciprocals,
2. negatives,
3. products,
4. quotients,
5. sums,
6. differences,

and otherwise to perform computations left to right. Bracketed quantities are treated as single numbers. In particular then, $x + yz$ means $x + (yz)$. Also, x/yz means $x/(yz)$, but $x/y + z$ means $(x/y) + z$. However, when quotients are indicated with a horizontal line, this line acts as brackets for what is above and below. So,

$$\frac{x + y}{z - w}$$

means $(x + y)/(z - w)$.

As a set with its operations $+$, $\cdot$ and $-$ and its particular elements 0 and 1, the real numbers can be denoted by $\mathbf{R}$. The algebraic properties of $\mathbf{R}$ are the properties of what is called a *field*. Two other fields are the set $\mathbf{Q}$ of *rational numbers* and the set $\mathbf{C}$ of *complex numbers*.

Definition. A **field** is a set containing distinct elements 0 and 1, and equipped with binary operations $+$ and $\cdot$ and a unary operation $-$, such that for all elements x, y and z, the following axioms are satisfied (when interpreted according to the notational conventions described above):

1. $x + (y + z) = x + y + z$;
2. $x + 0 = x$;
3. $x + (-x) = 0$;

4. $y + x = x + y$;
5. $x(y + z) = xy + xz$;
6. $(x + y)z = xz + yz$;
7. $x(yz) = xyz$;
8. $1x = x$;
9. if $x \neq 0$, then there exists w such that $xw = 1$;
10. $yx = xy$.

In Axiom 9, for each x different from 0, the element w is unique and can be denoted by x^{-1}. (*Proof:* if $xw = 1$ and $xv = 1$, then $xvw = 1w$, so (by Axioms 7 and 8) $x(vw) = w$, hence (by Axiom 10) $x(wv) = w$, hence $xwv = w$, hence $1v = w$, so finally $v = w$.)

Subtraction and division can be defined on any field as they are on **R**.

Axiom 6 follows from Axioms 5 and 10. The following facts are straightforward consequences of the axioms. For all x, y and z in a field, we have

- if $x + y = x + z$, then $y = z$;
- $0x = 0$;
- $-x = (-1)x$;
- if $xy = 0$, then $x = 0$ or $y = 0$.

Note that in any field, the elements 0 and 1 are combined according to the following tables:

$+$	0	1
0	0	1
1	1	

and

$\cdot$	0	1
0	0	0
1	0	1

.

Only the value of $1 + 1$ is not determined by the properties of a field. There is a field $\mathbf{F}_2$, whose only elements are 0 and 1 and in which $1 + 1 = 0$.

In a field **F**, the operation $+$, for example, can also be described as a *map* or *function* from $\mathbf{F} \times \mathbf{F}$ to **F**.

1 Basic properties of a vector space

Let **F** be a field. If you like, you can think throughout that **F** is the field **R** of real numbers. In any case, we may refer to **F** as the *scalar field*, and to its elements as *scalars*.

If n is a positive integer, we write $\mathbf{F}^n$ for the set of *ordered n-tuples*

$$(x_1, \dots, x_n)$$

of elements x_i of $\mathbf{F}$. We may abbreviate such an n-tuple by $\mathbf{x}$. Addition of such n-tuples can be defined in an obvious way:

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \dots, x_n + y_n).$$

Likewise, we can define $-\mathbf{x}$ to be $(-x_1, \dots, -x_n)$. We also write $\mathbf{0}$ for $(0, \dots, 0)$. With these definitions, Axioms 1, 2, 3 and 4 for fields hold.

It is possible to define a multiplication of n-tuples:

$$\mathbf{x} \cdot \mathbf{y} = (x_1 y_1, \dots, x_n y_n).$$

Equipped with these operations of addition and multiplication, and with the tuples $(0, \dots, 0)$ and $(1, \dots, 1)$ replacing 0 and 1 respectively, $\mathbf{F}^n$ satisfies the axioms of a field, *except* Axiom 9. Note also for example that in $\mathbf{R}^2$ we have $(1, 0) \cdot (0, 1) = (0, 0)$, but neither $(1, 0)$ nor $(0, 1)$ is equal to $(0, 0)$.

In linear algebra, we shall not be interested in the sort of multiplication just defined. Note however that we can also multiply a tuple by a scalar to get another tuple:

$$a\mathbf{x} = (ax_1, \dots, ax_n).$$

The operation so defined is called *scalar multiplication* (that is, multiplication by a scalar).

In the order of operations on $\mathbf{F}^n$, we shall treat scalar multiplication as we do multiplication of scalars: so, $a\mathbf{x} + \mathbf{y}$ means $(a\mathbf{x}) + \mathbf{y}$, and $\mathbf{x} + a\mathbf{y}$ means $\mathbf{x} + (a\mathbf{y})$. (A sum like $\mathbf{x} + a$ is not well-defined anyway.)

Equipped $+$, $-$, scalar multiplication and the element $\mathbf{0}$, the set $\mathbf{F}^n$ satisfies the following.

Definition. A **vector space** or **linear space** over the field $\mathbf{F}$ is a set V containing an element $\mathbf{0}$, equipped with a binary operation $+$ and a unary operation $-$, and equipped also, for each a in $\mathbf{F}$, with a unary operation $a\cdot$ taking $\mathbf{x}$ to $a\mathbf{x}$ (also written $a \cdot \mathbf{x}$), such that for each $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ in V, and for each a and b in $\mathbf{F}$, the following axioms hold:

1. $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = \mathbf{x} + \mathbf{y} + \mathbf{z}$;
2. $\mathbf{x} + 0 = \mathbf{x}$;
3. $\mathbf{x} + (-\mathbf{x}) = 0$;
4. $\mathbf{y} + \mathbf{x} = \mathbf{x} + \mathbf{y}$;
5. $a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$;
6. $(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$;
7. $a(b\mathbf{x}) = ab\mathbf{x}$;
8. $1\mathbf{x} = \mathbf{x}$.

Elements of a vector space are called **vectors**. Note the formal similarity of the vector-space axioms to the first eight axioms for fields. Note however that in vector spaces, the operation $\cdot$ combines things of different sorts (a scalar and a vector). We may write $\mathbf{x} - \mathbf{y}$ for $\mathbf{x} + (-\mathbf{y})$.

The following are consequences of the axioms, by the same proofs, formally, as the corresponding facts about fields. For all vectors $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ and for any scalar a, we have

- if $\mathbf{x} + \mathbf{y} = \mathbf{x} + \mathbf{z}$, then $\mathbf{y} = \mathbf{z}$;
- $0\mathbf{x} = \mathbf{0}$;
- $-\mathbf{x} = (-1)\mathbf{x}$;
- if $a\mathbf{x} = 0$, then $a = 0$ or $\mathbf{x} = 0$.

In particular, the operation $-$ and the element $\mathbf{0}$ of a vector space can be derived from the operation $\cdot$ and an arbitrary element of the space.

Note well that, strictly, a vector space is not a set, but a set with certain operations. Formally, one can understand a vector space to be a certain sort of sextuple (6-tuple), such as

$$(V, \mathbf{F}, +, \cdot, -, \mathbf{0}),$$

where the six terms are related as in the definition above. Again though, since $-$ and $\mathbf{0}$ can be derived from the rest of the terms, we can consider a vector space to be a quadruple (4-tuple), such as

$$(V, \mathbf{F}, +, \cdot).$$

For any such a quadruple to be a vector space, first, the following must hold:

- V must be a nonempty set;
- $\mathbf{F}$ must be a field;
- $+$ must be a binary operation on V, that is, a map from $V \times V$ to V;
- $\cdot$ must be a map $\mathbf{F} \times V \to V$.

If these hold, then the quadruple is the sort of structure for which the axioms in the definition make sense. If the axioms are *true* for the quadruple, then it is a vector space.

If $(V, \mathbf{F}, +, \cdot)$ is a vector space, we customarily let the set V stand for the whole space if there is no uncertainty about what the scalar field $\mathbf{F}$ is, or what the operations of addition or scalar multiplication are. Strictly, V is the *universe* or *underlying set* of the space.

2 Subspaces

Let $(V, \mathbf{F}, +, \cdot)$ be a vector space. A **subspace** of V is a subset W of V such that:

- $\mathbf{0}$ is in W;
- $\mathbf{x} + \mathbf{y}$ and $a\mathbf{x}$ are in W whenever $\mathbf{x}$ and $\mathbf{y}$ are in W and a is in $\mathbf{F}$.

In other words, a subspace of a vector space is a subset that is closed under the vector-space operations, including the operation distinguishing the zero-element. But this zero-element is $0 \cdot \mathbf{x}$ for any $\mathbf{x}$ in the space. Therefore, a subspace of a vector space is precisely a *nonempty* subset of the space that is closed under addition and scalar multiplication.

We can now recognize two sorts of problems:

1. Is a given quadruple $(V, \mathbf{F}, +, \cdot)$ a vector space?
2. Is a given subset of a known vector space $(V, \mathbf{F}, +, \cdot)$ a subset of V?

Note that a subspace is in fact a vector space itself, when equipped with the operations from the larger space. So, to establish that a particular quadruple is a vector space, it may be enough to show that it is a subspace of some known vector space.