LINEAR ALGEBRA - LECTURE 1

These are (somewhat expanded) notes on the first lecture on linear algebra, given the last day of fall semester. They are provided as a refresher for students who were in the fall semester class and a resource for students who, for example, took Honors Math A in 2017 and are now taking Honors Math B.

1. Definitions

Recall that in the process of writing down axioms for the real numbers, we defined the following concept:

**Definition 1.1.** A field is a set $F$ together with two binary operations $+$ (addition) and $\cdot$ (multiplication) and two distinguished elements 0 and 1, such that the following axioms hold:

- **Commutativity:** for all $x, y \in F$, $x + y = y + x$ and $x \cdot y = y \cdot x$.
- **Associativity:** for all $x, y, z \in F$, $(x + y) + z = x + (y + z)$ and $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.
- **Distributivity:** for all $x, y, z \in F$, $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$.
- **Identity axioms:** for all $x \in F$, $x + 0 = x$ and $x \cdot 1 = x$.
- **Existence of additive inverses:** for all $x \in F$, there exists a $y \in F$ such that $x + y = 0$. We denote this element $y$ by $-x$.
- **Existence of multiplicative inverses:** for all $x \in F$ such that $x \neq 0$, there exists a $y \in F$ such that $x \cdot y = 1$. We denote this element $y$ by $1/x$ or $x^{-1}$.

Examples of fields that we have seen so far are $\mathbb{Q}$, the rational numbers, and $\mathbb{R}$, the real numbers. At some point in the near future we will also define $\mathbb{C}$, the complex numbers.

**Example 1.2.** As a further example, let $p$ be a prime number and let $F_p = \{0, 1, \ldots, p - 1\}$ as a set. Define addition and multiplication by taking the “usual” sum or product and reducing modulo $p$ (i.e., taking the remainder of the result when you try to divide by $p$). Let 0 and 1 be the obvious choices. Then it is an exercise to show that $F_p$ is a field. As $F_p$ has finitely many elements, we say that it is a finite field.

Almost all of “high school algebra” still holds in an arbitrary field (keeping in mind that we are not talking about ordered fields at the moment, so inequalities are meaningless). An exception is that one has to be slightly careful about dividing by zero: there are fields, like $F_p$ above, for which $1 + 1 + \cdots + 1 = 0$, where we take $p$ 1’s in the sum. (Such fields are called characteristic $p$ fields; fields for which finite sums of 1 are never zero like $\mathbb{Q}$ and $\mathbb{R}$ and $\mathbb{C}$ are called characteristic 0 fields.) Therefore if we, entirely reasonably, denote this sum by $p$, we find that $p = 0$, so we cannot divide by $p$. Fortunately we will never need to do that in this class.

We will do linear algebra over an arbitrary field in the sense that we will prove theorems about vector spaces over arbitrary fields, not just real vector spaces (vector spaces over $\mathbb{R}$). Apostol only really discusses vector spaces over the fields $\mathbb{R}$ and $\mathbb{C}$, which do present special features that we will come to in time. We will be more general at the moment because it will be no more difficult: all of our proofs will go through in this greater generality in exactly the same way. This flexibility is one of the main strengths of linear algebra!

With this in mind, you should mostly feel free to think of all vector spaces as real vector spaces, if it is psychologically helpful.

Now for the main definition:

**Definition 1.3.** Let $F$ be a field. A vector space over $F$, or $F$-vector space, is a set $V$ together operations $+: V \times V \to V$ (vector addition) and $\cdot: F \times V \to V$ (scalar multiplication), as well as a distinguished element $0 \in V$, such that the following axioms hold:

- **Commutativity:** for all $X, Y \in V$, $X + Y = Y + X$.
- **Associativity:** for all $X, Y, Z \in V$, $(X + Y) + Z = X + (Y + Z)$, and for all $c, d \in F$ and $X \in V$ we have $c \cdot (d \cdot X) = (c \cdot d) \cdot X$.
- **Distributivity:** for all $c, d \in F$ and $X, Y \in V$, $c \cdot (X + Y) = (c \cdot X) + (c \cdot Y)$ and for all $c, d \in F$ and $X \in V$, $(c + d) \cdot X = (c \cdot X) + (d \cdot Y)$.
- **Identity axioms:** for all $X \in V$, $0 + X = X$ and $1 \cdot X = X$.
- **Inverses:** for all $X \in V$, there exists a $Y \in V$ such that $X + Y = 0$.
We often call an element of \( V \) a vector or point and an element of the base field \( F \) a scalar.

Note that we are overloading notation here: in the second associativity axiom, when we write \( d \cdot X \) we mean scalar multiplication in \( V \) and when we write \( c \cdot d \) we mean multiplication in \( F \). Similarly in the second distributivity axiom \( X + Y \) refers to vector addition while \( c + d \) refers to addition in \( F \). This notational abuse is common and we will usually not mention it from now on. As usual, we often omit the sign \( \cdot \) when writing scalar multiplication.

For reasons unknown to me, Apostol calls a vector space a linear space. Since I know of no one else who does this, please ignore him on this point.

**Proposition 1.4.** In any \( F \)-vector space \( V \), the following hold:

- \( 0 \) is unique.
- Given any \( X \in V \), \((-1)X\) is the unique \( Y \) such that \( X + Y = 0 \).
- For all \( X \in V \), \( 0X = 0 \).
- For all \( c \in F \), \( c0 = 0 \).
- For all \( c \in F \) and \( X \in V \), \((-c)X = c(-X) = -(cX)\).
- If \( cX = 0 \), then \( c = 0 \) or \( X = 0 \).
- If \( cX = cY \), then \( c = 0 \) or \( X = Y \).
- If \( cX = dY \), then \( c = d \) or \( X = 0 \).

**Proof:** Apply the definitions; all are straightforward. \( \square \)

We usually write \((-1)X\) as \(-X\).

### 2. Examples

**Example 2.1.** The most concrete vector spaces are Cartesian powers of the base field: if \( F \) is a field and \( n > 0 \) an integer then the set

\[
F^n = \{ \text{ordered } n\text{-tuples } (x_1, x_2, \ldots, x_n) \text{ with each } x_i \in F \}
\]

is naturally a vector space\(^1\) when we make the following definitions: we let \( 0 \) be the element \( (0, 0, \ldots, 0) \in F^n \), we define vector addition termwise like

\[
(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n),
\]

and we define scalar multiplication termwise like

\[
c \cdot (x_1, \ldots, x_n) = (cx_1, \ldots, cx_n).
\]

One then easily verifies that \( F^n \) is a vector space. In particular, if \( F = \mathbb{R} \) and \( n = 2 \), we find that the Cartesian plane is naturally a vector space; for more general \( n \) we find that Euclidean space \( \mathbb{R}^n \) is a vector space as well.

There is some convenient terminology associated with these vector spaces: if \( x \) is a vector in \( F^n \), then by definition we can write it as

\[
x = (x_1, \ldots, x_n)
\]

with each \( x_i \in F \). We call \( x_i \) the \( i \)th coordinate or \( i \)th component of \( x \). Thus two vectors in \( F^n \) are equal if and only if all of their coordinates are equal, and the zero vector is the vector with all coordinates equal to zero.

There is a useful degenerate case which we will define now: \( F^0 \) is the trivial vector space over \( F \), with only one element \( \{0\} \) and the obvious operations. It’s not a very interesting vector space, but it’s sometimes convenient to have around. Note also that \( F^1 \) is, as a set, simply \( F \) itself, and the scalar multiplication and vector addition operations are the obvious ones induced from \( F \). It is useful to distinguish the field \( F \) from the vector space \( F^1 \), even though they look like “the same” object; in some sense \( F^1 \) “forgets” which element is the multiplicative identity \( 1 \) while the field \( F \) remembers that data. This distinction will not be important in this course.

For low values of \( n \) (up to about 3) we can also “draw” these vector spaces, especially if \( F = \mathbb{R} \), just by drawing the Cartesian plane or 3-space; vectors just correspond to points in the space. When we are thinking of the points as vectors, we usually draw them by drawing a straight arrow from the origin (zero vector) to the point in question. Scalar multiplication then “scales” the arrow in question by the relevant amount (and flips the arrow about the origin if the amount is negative). Vector addition can be visualized

---

\(^1\)Recall that the an ordered \( n \)-tuple of elements of \( F \) is simply a function \( x : \{1, \ldots, n\} \to F \), and we generally write \( x_i \) in place of \( x(i) \).
via the “parallelogram rule”: if \( \mathbf{x} \) and \( \mathbf{y} \) are two vectors, then if we draw the parallelogram which has the arrows corresponding to \( \mathbf{x} \) and \( \mathbf{y} \) as two of its sides the vector \( \mathbf{x} + \mathbf{y} \) is the point represented by the corner opposite to the origin.

**Example 2.2.** Another source of vector spaces is as spaces of functions: if \( S \) is any set and \( F \) is a field, then let \( \mathcal{F}(S, F) \) denote the set of functions \( f : S \to \mathbb{R} \). Define addition by \((f + g)(s) = f(s) + g(s)\), scalar multiplication by \((c \cdot f)(x) = c \cdot f(x)\), and let \( 0 \) be the function taking all elements in \( S \) to 0 \( \in F \). Then \( \mathcal{F}(S, F) \) is a vector space over \( F \).

Just for practice, let’s carefully verify the first couple of field axioms; the only difficulty comes in figuring out exactly what the notations mean in every instance. For commutativity, we note that if \( f, g \in \mathcal{F}(S, F) \), then for each \( s \in S \) we have

\[
(f + g)(s) = f(s) + g(s) = g(s) + f(s) = (g + f)(s).
\]

Therefore \( f + g = g + f \) in \( \mathcal{F}(S, F) \). For the first and third equality we used the definition of vector addition in our vector space, and for the second equality we used the commutativity of addition in \( F \). Note how the symbol “+” is being used for two different things: it refers to addition in \( F \) in the second and third lines and vector addition in \( \mathcal{F}(S, F) \) in the first and fourth. For the first associativity axiom, we note that if \( f, g, h \in \mathcal{F}(S, F) \), then for each \( s \in S \) we have

\[
((f + g) + h)(s) = ((f + g)(s)) + h(s)
= (f(s) + g(s)) + h(s)
= f(s) + (g + h)(s)
= f(s) + (g + h)(s)
= (f + (g + h))(s).
\]

On the first, second, fourth, and fifth lines we have simply used the definition of vector addition in \( \mathcal{F}(S, F) \), while in the third line we used associativity of addition in \( F \). For the second associativity axiom, we have for every \( c, d \in F \) and every \( f \in \mathcal{F}(S, F) \) that

\[
(c \cdot (d \cdot f))(s) = c \cdot (d \cdot f)(s) = c \cdot (d \cdot f(s)) = (c \cdot d) \cdot f(s).
\]

I leave it as an exercise to determine exactly what facts we are using in the calculation above, as well as the verification of the other vector space axioms.

**Example 2.3.** We can generalize the previous example: suppose that we have already an \( F \)-vector space \( V \). Then we can equip the set of functions \( f : S \to V \), denoted \( \mathcal{F}(S, V) \), with the structure of an \( F \)-vector space in exactly the same way as above: addition and scalar multiplication are done pointwise, and the zero function is the zero vector. Now it is the axioms of the vector space \( V \) that ensure that \( \mathcal{F}(S, V) \) will be a vector space as well.

### 3. Subspaces

One way to produce more vector spaces is to start with a given vector space and look “inside” it.

**Definition 3.1.** Let \( V \) be an \( F \)-vector space. A nonempty subset \( W \subseteq V \) is a **subspace** if for all \( \mathbf{X}, \mathbf{Y} \in V \) and \( c \in F \), \( \mathbf{X} + \mathbf{Y} \in W \) and \( \mathbf{X} \in W \) implies that \( \mathbf{X} + \mathbf{Y} \in W \) and \( \mathbf{X} \in W \) implies that \( c \mathbf{X} \in W \).

**Proposition 3.2.** If \( W \) is a subspace of an \( F \)-vector space \( V \), then \( W \) is itself an \( F \)-vector space (with scalar multiplication, vector multiplication, and zero vector inherited from \( V \)).

The proof is essentially immediate: the definition of a subspace ensures that + and \( \cdot \) are well-defined, and the axioms of \( V \) imply the corresponding axioms for \( W \). A subspace is nothing other than a subset that also happens to be a vector space when the operations are appropriately restricted.

**Example 3.3.** Given any \( F \)-vector space \( V \), the trivial subspace \( W = \{0\} \) consisting only of the zero vector is always a subspace. On the opposite extreme, \( V \) itself is a subspace of \( V \).

**Example 3.4.** Any line through the origin in \( \mathbb{R}^n \) is a subspace. How can you see this geometrically? A line that does not pass through the origin in \( \mathbb{R}^n \) is never a subspace of \( \mathbb{R}^n \).

**Example 3.5.** We can generalize the previous example as follows: a line through the origin in \( \mathbb{R}^n \) can be described as \( \{c\mathbf{x} \mid c \in \mathbb{R}\} \), where \( \mathbf{x} \) is any fixed nonzero point on the line. The same procedure works in general: if \( V \) is an \( F \)-vector space and \( \mathbf{X} \) an element of \( V \), then \( \{c\mathbf{X} \mid c \in F\} \) is a subspace of \( V \) (if \( \mathbf{X} = 0 \), it is the trivial subspace).
Example 3.6. Many conditions in calculus give rise to natural subspaces of the space of functions from $\mathbb{R}$ to $\mathbb{R}$, which we saw above is a vector space $F(\mathbb{R}, \mathbb{R})$. For example, let $C(\mathbb{R}, \mathbb{R})$ be the set of continuous functions $\mathbb{R} \to \mathbb{R}$. We know that if $f$ and $g$ are continuous, then $f + g$ and $cf$ are continuous (for any $c \in \mathbb{R}$), so by definition $C(\mathbb{R}, \mathbb{R})$ is a subspace of $F(\mathbb{R}, \mathbb{R})$. Other examples of subspaces include the space of bounded functions, the space of differentiable functions, the space of five times differentiable functions, the space of smooth functions, the space of analytic functions...

Example 3.7. Considering instead $F([a, b], \mathbb{R})$, where $a < b$ are real numbers, the set of integrable functions $[a, b] \to \mathbb{R}$ is a subspace.

Example 3.8. We can also consider vector spaces of sequences of real numbers: recall that a sequence is simply a function $\mathbb{Z}_{>0} \to \mathbb{R}$, so sequences form a vector space $F(\mathbb{Z}_{>0}, \mathbb{R})$. Here is a chain of natural subspaces:

$$
\begin{align*}
F(\mathbb{Z}_{>0}, \mathbb{R}) & \subseteq \{\text{bounded sequences}\} \\
& \subseteq \{\text{convergent sequences}\} \\
& \subseteq \{\text{sequences that converge to 0}\} \\
& \subseteq \{\text{sequences whose corresponding series converges}\} \\
& \subseteq \{\text{sequences whose corresponding series converges to 0}\}
\end{align*}
$$

The fact that each is a vector space is a simple check of the definition, combined with facts from calculus. On the other hand, the set of sequences converging to 7 is not a subspace (why?).

Intersections of subspaces are automatically subspaces:

**Proposition 3.9.** Let $V$ be an $F$-vector space and $W_1, W_2$ subspaces of $V$. Then $W_1 \cap W_2$ is a subspace of $V$.

**Proof.** We need to check that if $X, Y \in W_1 \cap W_2$ and $c \in F$, then $X + Y \in W_1 \cap W_2$ and $cX \in W_1 \cap W_2$. By the definition of intersection, we find that $X \in W_1$ and $X \in W_2$, and similarly for $Y$. Since $W_1$ and $W_2$ are subspaces by assumption, we find that $X + Y$ lies in $W_1$, and also that $X + Y$ lies in $W_2$. By the definition of intersection, $X + Y \in W_1 \cap W_2$. Similarly, $cX \in W_1$ and $cX \in W_2$, so $cX \in W_1 \cap W_2$ as well. 

**Example 3.10.** Letting $V = F(\mathbb{R}, \mathbb{R})$, $W_1$ the subspace of continuous functions, and $W_2$ the space of bounded functions, the proposition tells us that the space $W_1 \cap W_2$ of bounded and continuous functions is a subspace.

Unions of subspaces are usually not subspaces: for example, the union of two different lines through the origin in $\mathbb{R}^2$ is very much not a subspace of $\mathbb{R}^2$. 