Lecture 1e

Vector Spaces

All of this leads us to notice that quite a lot of things that we frequently encounter have a common underlying structure. And so, instead of studying these things individually, we instead will study them in general, based only on this common structure. In a nod to $\mathbb{R}^n$, that started it all, they are known as vector spaces.

Definition: A **vector space over** $\mathbb{R}$ is a set $\mathbb{V}$ together with an operation of **addition**, usually denoted $\mathbf{x}+\mathbf{y}$ for any $\mathbf{x},\mathbf{y}\in\mathbb{V}$, and an operation of **scalar multiplication**, usually denoted $s\mathbf{x}$ for any $\mathbf{x}\in\mathbb{V}$ and $s\in\mathbb{R}$, such that for any $\mathbf{x},\mathbf{y},\mathbf{z}\in\mathbb{V}$ and $s,t\in\mathbb{R}$ we have all of the following properties:

(V1) $\mathbf{x}+\mathbf{y}\in\mathbb{V}$ (closed under addition)
(V2) $(\mathbf{x}+\mathbf{y})+\mathbf{z}=\mathbf{x}+(\mathbf{y}+\mathbf{z})$ (addition is associative)
(V3) There is an element $\mathbf{0}\in\mathbb{V}$, (called the **zero vector**), such that

$$\mathbf{x}+\mathbf{0}=\mathbf{x}=\mathbf{0}+\mathbf{x} \quad \text{(additive identity)}$$

(V4) For each $\mathbf{x}\in\mathbb{V}$, there exists an element $-\mathbf{x}$ such that $\mathbf{x}+(-\mathbf{x})=\mathbf{0}$. (additive inverse)
(V5) $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}$ (addition is commutative)
(V6) $s\mathbf{x}\in\mathbb{V}$ (closed under scalar multiplication)
(V7) $s(t\mathbf{x})=(st)\mathbf{x}$ (scalar multiplication is associative)
(V8) $(s+t)\mathbf{x}=s\mathbf{x}+t\mathbf{x}$ (scalar addition is distributive)
(V9) $s(\mathbf{x}+\mathbf{y})=s\mathbf{x}+s\mathbf{y}$ (scalar multiplication is distributive)
(V10) $1\mathbf{x}=\mathbf{x}$ (scalar multiplicative identity)

Note: In general, an element of a vector space $\mathbb{V}$ is known as a vector. As elements of $\mathbb{R}^n$ are also known as vectors, this can be confusing, so to help we use the notation $\mathbf{x}$ to mean a vector from a general vector space, and reserve the symbol $\vec{x}$ to mean an element of $\mathbb{R}^n$.

We have already seen that $\mathbb{R}^n$ is a vector space, as well as the set of $m\times n$ matrices, and polynomials of degree up to n.

Notation: We write $M(m,n)$ for the vector space of $m\times n$ matrices, and we write P_n for the vector space of polynomials of degree up to n.

Some other vector spaces are:

$\mathcal{F}$: the set of all functions from $\mathbb{R}$ to $\mathbb{R}$; and

$\mathcal{F}(a,b)$: the set of all functions from the interval (a,b) to $\mathbb{R}$.

There are many more common vector spaces, which we will explore throughout the course, but first I'd like to take a look at a non-standard vector space.

Things get interesting when non-standard definitions of addition and scalar multiplication are used. In these cases, the usual notation for addition and scalar multiplication are replaced with the symbols $\oplus$ and $\odot$. (Sometimes $\oplus_{\mathbb{V}}$ and $\odot_{\mathbb{V}}$ are used if we need to keep track of which vector space we are referring to.) Let's look at this example, which is started but not completed in the text:

Example 10: Let $\mathbb{V} = \{(a,b) \mid a,b \in \mathbb{R}, b > 0\}$. We will define addition in $\mathbb{V}$ by $(a,b) \oplus (c,d) = (ad+bc, bd)$ and we define scalar multiplication in $\mathbb{V}$ by $t \odot (a,b) = (tab^{t-1}, b^t)$. Now let's show that $\mathbb{V}$ is a vector space, paying close attention to how the axioms look with our unusual definitions. To that end, let $(a,b),(c,d),(e,f) \in \mathbb{V}$, and let $s,t \in \mathbb{R}$.

(V1): $(a,b) \oplus (c,d) = (ad+bc, bd)$, where $ad+bc \in \mathbb{R}$ and $bd \in \mathbb{R}$, and since both $b > 0$ and $d > 0$, we have that $bd > 0$. This means that $(ad+bc, bd) \in \mathbb{V}$, and thus $(a,b) \oplus (c,d) \in \mathbb{V}$.

(V2): $((a,b)\oplus(c,d))\oplus(e,f) = (ad+bc, bd)\oplus(e,f) = ((ad+bc)f+(bd)e, (bd)f) = (adf+bcf+bde, bdf) = (a(df)+b(cf+de), b(df)) = (a,b) \oplus (cf+de, df) = (a,b)\oplus((c,d)\oplus(e,f))$ Note that, thanks to property V1, we don't need to worry about whether or not any of the intermediate steps are in $\mathbb{V}$.

(V3): To prove this property, we need to find an element $\mathbf{0}$ of $\mathbb{V}$ such that $(a,b) \oplus \mathbf{0} = (a,b)$. Let's assume that $\mathbf{0} = (x,y)$. Then we want $(a,b) \oplus (x,y) = (ay-bx, by) = (a,b)$. So we want $b = by$, which means $y = 1$, and then we want $ay - bx = a$, and plugging in $y = 1$, this becomes $a - bx = a$, so $-bx = 0$. Now $b \neq 0$, since $b > 0$, so the only way to have $-bx = 0$ is to have $x = 0$. All this work leads us to the GUESS that $\mathbf{0} = (0,1)$. Now we need to prove it. First we note that $(0,1) \in \mathbb{V}$, since $0,1 \in \mathbb{R}$ and $1 > 0$. Next, we note that $(a,b) \oplus (0,1) = ((a)(1)+(b)(0), (b)(1)) = (a,b)$ and $(0,1) \oplus (a,b) = ((0)(b)+(1)(a), (1)(b)) = (a,b)$. And so we see that V3 holds, with $(0,1)$ as our zero vector.

(V4): We found in V3 that $\mathbf{0} = (0,1)$. Now, given (a,b), we need to find $-(a,b) \in \mathbb{V}$. So let's look for (w,z) such that $(a,b) \oplus (w,z) = (0,1)$. Well, $(a,b) \oplus (w,z) = (az+bw, bz)$, so we need $bz = 1$ and $az + bw = 0$. From $bz = 1$ we get that $z = (1/b)$. (Note that we can divide by b, since $b > 0$.) Plugging this into $az + bw = 0$, we get $a(1/b) + bw = 0$, so $w = (-a/b^2)$. So we now guess that $-(a,b) = (-a/b^2, 1/b)$. First, we note that since $b > 0$, $1/b > 0$, so $(-a/b^2, 1/b) \in \mathbb{V}$. Next, we see that $(a,b)\oplus(-a/b^2, 1/b) = (a(1/b) + b(-a/b^2), b(1/b)) = (a/b - a/b, 1) = (0,1)$, as desired.

(V5): $(a,b) \oplus (c,d) = (ad+bc, bd) = (cb+da, db) = (c,d) \oplus (a,b)$

(V6): $s \odot (a,b) = (sab^{s-1}, b^s)$. Since $b > 0$, $b^s > 0$ for any s, and of course $sab^{s-1}, b^s \in \mathbb{R}$, so $(sab^{s-1}, b^s) \in \mathbb{V}$, which means $s \odot (a,b) \in \mathbb{V}$.

(V7): $s\odot(t\odot(a,b)) = s\odot((tab^{t-1}, b^t) = (stab^{t-1}(b^t)^{s-1}, (b^t)^s) = (stab^{t-1}b^{ts-t}, b^{ts}) = (stab^{t-1+ts-t}, b^{ts}) = (stab^{ts-1}, b^{ts}) = (stab^{st-1}, b^{st}) = (st) \odot (a,b)$

(V8): $(s+t)\odot(a,b) = ((s+t)ab^{s+t-1}, b^{s+t}) = (sab^{s-1}b^t + tab^{t-1}b^s, b^s b^t) = (sab^{s-1}, b^s) \oplus (tab^{t-1}, b^t) = (s\odot(a,b)) \oplus (t\odot(a,b))$

(V9): $s\odot((a,b)\oplus(c,d)) = s\odot(ad+bc, bd) = (s(ad+bc)(bd)^{s-1}, (bd)^s) = (sad(bd)^{s-1} + sbc(bd)^{s-1}, b^s d^s) = (sab^{s-1}d^s + scd^{s-1}b^s, b^s d^s) = (sab^{s-1}, b^s) \oplus (scd^{s-1}, d^s) = (s\odot(a,b)) \oplus (s\odot(c,d))$.

(V10): $1\odot(a,b) = (1ab^{1-1}, b^1) = (ab^0, b) = (a(1), b) = (a,b)$.

Examples such as this one force us to really pay attention to what the vector properties are saying. Thankfully, such situations rarely occur, as we will mostly focus on the standard vector spaces. With the standard vector spaces, our properties are obviously true. So we will not focus our attention so much on proving that these properties hold, but instead will use the properties (known to be true for all vector spaces) to prove other facts that will also be true for all vector spaces. The first example of this is the following theorem:

<u>Theorem 4.2.1</u>: Let $\mathbb{V}$ be a vector space. Then
(1) $0\mathbf{x} = \mathbf{0}$ for all $\mathbf{x} \in \mathbb{V}$
(2) $(-1)\mathbf{x} = -\mathbf{x}$ for all $\mathbf{x} \in \mathbb{V}$
(3) $t\mathbf{0} = \mathbf{0}$ for all $t \in \mathbb{R}$

I'll prove property (2) now (leaving the proof of (3) as a practice problem):

<u>Proof of Theorem 4.2.1(2)</u>: Before I actually dive into the proof, I want to talk about what it says. Because, up until now, we have been taking it as a notational convention that $(-1)\mathbf{x} = -\mathbf{x}$. But in this section we introduce the notation $-\mathbf{x}$ to mean the additive inverse of $\mathbf{x}$, and not necessarily the scalar product of -1 times $\mathbf{x}$. Of course, the point of this theorem is to show that these two values are in fact equal, thus justifying our earlier decision to set them equal.

Now, to start the proof, I'll actually want to prove another fact first: that the additive inverse is unique. That is to say, if $\mathbf{x}+\mathbf{y} = \mathbf{0}$ and $\mathbf{x}+\mathbf{z} = \mathbf{0}$, then $\mathbf{y} = \mathbf{z}$. To see this, let $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ be as stated, and notice that

$$\begin{aligned}
\mathbf{z} &= \mathbf{0} + \mathbf{z} && \text{by V3}\\
&= (\mathbf{x}+\mathbf{y}) + \mathbf{z} && \text{by our choice of } \mathbf{y}\\
&= (\mathbf{y}+\mathbf{x}) + \mathbf{z} && \text{by V5}\\
&= \mathbf{y} + (\mathbf{x}+\mathbf{z}) && \text{by V2}\\
&= \mathbf{y} + \mathbf{0} && \text{by our choice of } \mathbf{z}\\
&= \mathbf{y} && \text{by V3}
\end{aligned}$$

Thanks to the uniqueness of the additive inverse, we now know that in order to show that some $\mathbf{y}$ equals the additive inverse of $\mathbf{x}$ (i.e., to show $\mathbf{y} = -\mathbf{x}$), we need to show that it satisfies the condition in property V4: $\mathbf{x}+(-\mathbf{x}) = \mathbf{0}$. In our particular case, we suspect that $-\mathbf{x}$ is $(-1)\mathbf{x}$, and so we will look at $\mathbf{x}+(-1)\mathbf{x}$:

$$
\begin{aligned}
\mathbf{x} + (-1)\mathbf{x} &= (1)\mathbf{x} + (-1)\mathbf{x} && \text{by V10} \\
&= (1 + (-1))\mathbf{x} && \text{by V8} \\
&= (0)\mathbf{x} && \text{operation of numbers in } \mathbb{R} \\
&= \mathbf{0} && \text{by Theorem 4.2.1(1)}
\end{aligned}
$$

And since $\mathbf{x} + (-1)\mathbf{x} = \mathbf{0}$, we know that $(-1)\mathbf{x} = -\mathbf{x}$.