Chapter 24

Vector Spaces

24.1 Definition of a vector space

We start with describing the abstract definition of a vector space. The idea is that vectors do not need to be just the n-tuples in $\mathbb{R}^n$ that we are used to, but instead, we abstract the critical properties that vectors in $\mathbb{R}^n$ have and then any other set of objects that satisfy these properties will also form a vector space. If we can prove properties of the vector space using just these bare requirements, then anything that satisfies these properties also satisfies the theorems we prove. This is in essence how mathematicians use abstraction to generalize properties. For physics, these ideas become most important when we study quantum mechanics. Quantum mechanics works with the abstraction that functions are infinite-dimensional vectors. We won't go into extensive detail here, but this is why these ideas and the ability to use abstraction are so important to physicists.

Vector space: A vector space is a collection of objects $v \in V$ that have the arithmetic operations of addition and multiplication defined by the following postulates:

1. If $\vec{u}$ and $\vec{v}$ are in V, then so is $\vec{u} + \vec{v}$ (closure under addition)

2. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ (vector addition is commutative)

3. $u + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$ (vector addition is associative)

4. There is a zero vector such that $\vec{v} + \vec{0} = \vec{v}$ for all $v \in V$. (existence of an additive identity)

5. If $\vec{v}$ is in V, then there is a vector $-\vec{v}$ such that $\vec{v} + (-\vec{v}) = \vec{0}$. This is called the negative of $\vec{v}$. (existence of an additive inverse)
6. If a is a scalar and $\vec{v} \in V$, then $a\vec{v} \in V$ (closure under scalar multiplication)
7. $a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$ (distributive property)
8. $(a + b)\vec{v} = a\vec{v} + b\vec{v}$ (distributive property)
9. $(ab)\vec{v} = a(b\vec{v})$ (scalar multiplication is associative)
10. $1\vec{v} = \vec{v}$ (existence of a scalar multiplicative identity)

If all scalars are real numbers, this is a *real vector space*. If all scalars are complex numbers, this is a *complex vector space*. Note this definition comes from the scalars in the vector space. Ordinary vectors like $\mathbb{R}^3$ are real vector spaces.

A subspace is a vector space $S \subset V$ such that all vectors in the subspace form a vector space. An example of this is that $\mathbb{R}^2$ is a subset of $\mathbb{R}^3$. Note: The zero vector is always in a vector subspace because any vector space must have an additive inverse.

A set of vectors $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \ldots, \vec{v}_n, \}$ are *dependent* if there exist scalars $\alpha_1, \ldots, \alpha_n$ such that $\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n = 0$ and not all scalars α_i are zero. This definition can be difficult to understand. Let's look at an example: $(1, 0, 0), (1, 1, 0), (0, 1, 0)$ are dependent since $-1 \times (1, 0, 0) + 1 \times (1, 1, 0) + (-1) \times (0, 1, 0) = 0$. Note how at least one of the scalars is nonzero—one does not need all of them to be nonzero, usually at least two is sufficient.

A set $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \ldots, \vec{v}_n\}$ is said to span V if every vector in V can be written as $\vec{v} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n$. If the set of vectors is dependent, we can reduce it by removing basis vectors that can be expressed in terms of the remaining ones until we finally reach a set of independent spanning vectors. The *number* of these vectors is called the *dimension* of V. Because these vectors are independent, the scalars used to represent any vector are unique to that vector. They are called the coordinates of the vector in the given basis. While it may sound abstract, you are familiar with it. Consider the basis vectors $\hat{i} = (1, 0, 0)$, $\hat{j} = (0, 1, 0)$, and $\hat{k} = (0, 0, 1)$ in $\mathbb{R}^3$. The normal way we write a vector as $\vec{v} = (v_1, v_2, v_3)$ is actually a shorthand for $v_1\hat{i} + v_2\hat{j} + v_3\hat{k}$—we already know that these numbers v_1, v_2 and v_3 are a *unique* representation of the vector $\vec{v}$.

24.2 Scalar product

Let V be a complex vector space. Then the complex-valued scalar product satisfies:

1. $(\vec{u} \cdot \vec{v}) = (\vec{v} \cdot \vec{u})^*$ (note the complex conjugate!)

2. $\vec{u} \cdot (\vec{v} + \vec{w}) = (\vec{u} \cdot \vec{v}) + (\vec{u} \cdot \vec{w})$

3. $(a\vec{u} \cdot \vec{v}) = a(\vec{u} \cdot \vec{v})$ (Note that this means $\vec{u} \cdot (a\vec{v}) = a^*\vec{u} \cdot \vec{v}$)

4. $(\vec{u} \cdot \vec{u}) \geq 0$ for all $\vec{u} \in V$

5. $(\vec{u} \cdot \vec{u}) = 0$ implies $\vec{u} = \vec{0}$.

Example: n-tuples of complex numbers: $\vec{u} = (u_1, ...u_n)$, $\vec{v} = (v_1, ...v_n)$. So $(\vec{u} \cdot \vec{v}) = u_1 v_1^* + ... + u_n v_n^*$. You should verify that this satisfies all properties of a scalar product.

We now define a norm, which can be thought of as the length of a vector. A norm is less general than an inner product. We will see all vector spaces with inner products automatically have norms, but the converse is not always true—a normed vector space need not also have an inner product.

Norm: The norm is defined by $\|\vec{u}\| = (\vec{u} \cdot \vec{u})^{\frac{1}{2}}$

1. $\|\vec{u}\| \geq 0$

2. $\|\vec{u}\| = 0 \implies \vec{u} = \vec{0}$

3. $\|a\vec{u}\| = |a|\|\vec{u}\|$

4. $|(\vec{u} \cdot \vec{v})| \leq \|\vec{u}\|\|\vec{v}\|$, this is called the Cauchy Identity.

5. $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$, this is called the Triangle Inequality.

Proof of property 4 (Cauchy inequality): If $\vec{u} = 0$ or $\vec{v} = 0$, then property 4 obviously holds, so let's assume $\vec{u} \neq 0$ and $\vec{v} \neq 0$. Also assume that $(\vec{u} \cdot \vec{v}) \neq 0$, otherwise the inequality also obviously holds. So we define

$\lambda = \frac{|(\vec{u}\cdot\vec{v})|}{(\vec{u}\cdot\vec{v})}$, and note that $|\lambda| = 1$. Then consider

$$\begin{aligned}
0 \leq \left\| \frac{\lambda\vec{u}}{\|u\|} - \frac{\vec{v}}{\|v\|} \right\|^2 &= |\lambda|^2 \frac{\|u\|^2}{\|u\|^2} - \frac{\lambda(\vec{u}\cdot\vec{v}) + \lambda^*(\vec{v}\cdot\vec{u})}{\|u\|\|v\|} + \frac{\|v\|^2}{\|v\|^2} \\
&= 1 - \frac{|(\vec{u}\cdot\vec{v})| + |(\vec{v}\cdot\vec{u})|}{\|u\|\|v\|} + 1 \\
\implies 2|(\vec{u}\cdot\vec{v})| \leq 2\|u\|\|v\| & \qquad (24.1)
\end{aligned}$$

so, $|(\vec{u}\cdot\vec{v})| \leq \|u\|\|v\|$.

Now we prove the Triangle Inequality (Property 5):

$$\begin{aligned}
\|\vec{u}+\vec{v}\|^2 &= (\vec{u}+\vec{v})\cdot(\vec{u}+\vec{v}) \\
&= (\vec{u}\cdot\vec{u}) + (\vec{u}\cdot\vec{v}) + (\vec{v}\cdot\vec{u}) + (\vec{v}\cdot\vec{v}) \\
&= \|\vec{u}\|^2 + 2\,\mathrm{Re}\,(\vec{u}\cdot\vec{v}) + \|\vec{v}\|^2 \\
&\leq \|\vec{u}\|^2 + 2|(\vec{u}\cdot\vec{v})| + \|\vec{v}\|^2 \\
&\leq \|\vec{u}\|^2 + 2\|\vec{u}\|\|\vec{v}\| + \|\vec{v}\|^2 \\
&\leq (\|\vec{u}\| + \|\vec{v}\|)^2 \\
\implies \|\vec{u}+\vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|. & \qquad (24.2)
\end{aligned}$$

24.3 Polynomials and functions as vectors

Let's now try a nontrivial example. Let $V_n = \{$polynomials of degree $< n$ with complex coefficients$\}$ which is a complex vector space and let $(\vec{u}\cdot\vec{v}) = u_1\cdot v_1^* + ... + u_n\cdot v_n^*$. We will denote the two vectors as $\vec{u} = u_1 + u_2x + ... + u_nx^{n-1}$, and $\vec{v} = v_1 + v_2x + ... + v_nx^{n-1}$ (note the overarrow is schematic, these are *not* n-tuples in a conventional vector space). If we think of the powers of x in the polynomial as the "placeholder" for the different elements in a conventional n-dimensional vector, then we can immediately see that this vector space is isomorphic to $\mathbb{C}^n$. We can think of the basis vectors being $1, x, x^2, \cdots, x^{n-1}$, and so on.

A more complex example is the vector space of real-valued continuous functions on the interval $x : 0 \leq x \leq 1$. The scalar product is:

$$(f\cdot g) = \int_0^1 dx f(x)g(x). \qquad (24.3)$$

On can immediately verify that all properties of a scalar product hold—the only one that really needs to be checked is

$$(f \cdot f) = \int_0^1 dx \Big(f(x)\Big)^2 \geq 0, \tag{24.4}$$

which is clear for real-valued functions.

What is a basis for this vector space? We can show that $\{\sin(2\pi x), \sin(4\pi x), \sin(6\pi x), \cdots, \sin(2n\pi x)\}$ are all independent. If they were not independent, then

$$c_1 \sin(2\pi x) + c_2 \sin(4\pi x) + c_3 \sin(6\pi x) + ... + c_n \sin(2n\pi x) = 0 \tag{24.5}$$

for some set of $c_1 ... c_n$ that are not all zero. Multiply the above by $\sin(2m\pi x)$ and integrate from 0 to 1. Note that

$$\sin^2(2m\pi x) = \frac{[]1 - \cos(4m\pi x)]}{2} \tag{24.6}$$

and

$$\sin(2m\pi x)\sin(2m'\pi x) = \frac{[\cos(2m\pi - 2m'\pi)x - \cos(2m\pi + 2m'\pi)x]}{2} \tag{24.7}$$

for $m \neq m'$. So we can integrate to find

$$\int_0^1 dx \sin^2(2m\pi x) = \int_0^1 \frac{[]1 - \cos(4m\pi x)]}{2} = \frac{1}{2} - \frac{\sin(4m\pi x)}{8m\pi}\Bigg|_0^1 = \frac{1}{2}$$

$$\int_0^1 dx \sin(2m\pi x)\sin(2m'\pi x) = \int_0^1 dx \left[\frac{\cos(2m\pi x - 2m'\pi x) - \cos(2m\pi x + 2m'\pi x)}{2}\right]$$

$$= \frac{1}{2}\left[\frac{\sin(2m\pi x - 2m'\pi x)}{2m\pi - 2m'\pi} - \frac{\sin(2m\pi x + 2m'\pi x)}{2m\pi + 2m'\pi}\right]\Bigg|_0^1 = 0. \tag{24.8}$$

So we find that $c_m = 0$. We repeat this for all m and find the solution holds for all m. This then implies that $c_1 = c_2 = ... c_n = 0$, which implies this set of functions is independent for this vector space.