Math 54: Linear Algebra and Differential Equations

Jeffrey Kuan

July 19, 2019

Inner Product Spaces

Remember that linear algebra is the study of algebraic structure. The structure that we have been considering so far has been the structure of a vector space over $\mathbb{R}$. But in many ways, we see that this structure does not complete encompass all the properties we might want about a vector space. For example, consider $\mathbb{R}^2$. Even though $\mathcal{B}_1 = \{(1,0),(0,1)\}$ and $\mathcal{B}_2 = \{(1,1),(2,1)\}$ are both bases, in some sense, we prefer $\mathcal{B}_1$ because the vectors are at right angles to each other. In particular, $\mathbb{R}^n$ has an additional structure, where we can measure lengths and angles. So we will consider the structure of an inner product space over $\mathbb{R}$.

Definition: An **inner product space over** $\mathbb{R}$ is a vector space over $\mathbb{R}$ equipped with a bilinear form $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ called an **inner product** that satisfies the following properties.

- $\langle c_1 v_1 + c_2 v_2, w \rangle = c_1 \langle v_1, w \rangle + c_2 \langle v_2, w \rangle$
- $\langle v, c_1 w_1 + c_2 w_2 \rangle = c_1 \langle v, w_1 \rangle + c_2 \langle v, w_2 \rangle$ (bilinearity)
- $\langle v, v \rangle \geq 0$, where $\langle v, v \rangle = 0$ if and only if $v = 0$ (positive definite)
- $\langle v, w \rangle = \langle w, v \rangle$ (symmetry of the inner product)

Given an inner product space, we have a natural notion of length. We define for each vector v the **norm** of the vector $||v||$, which in some sense gives the "length" of the vector. The norm is defined by

$$||v|| = \langle v, v \rangle^{1/2}$$

We can take a square root on the right hand side by the positive definiteness of the inner product, which ensures that $\langle v, v \rangle \geq 0$. Using the norm, we can define the **distance between two vectors** u and v as $||u - v||$. Given any nonzero vector v, the vector of length 1 that points in the direction of v, defined as $\frac{v}{||v||}$ is called the **unit vector in the direction of** v.

Given an inner product space, we also have a natural notion of angle. The angle θ between two vectors is the θ value between 0 and π radians such that

$$\cos(\theta) = \frac{\langle v, w \rangle}{||v|| \cdot ||w||}$$

In particular, note that if $\theta = \frac{\pi}{2}$ (a right angle), then $\langle v, w \rangle = 0$. In particular, we say that two vectors u and v are **orthogonal** if and only if $\langle u, v \rangle = 0$.

For any inner product space, we have the following three fundamental properties. First, the **Triangle Inequality**:

$$||u + v|| \leq ||u|| + ||v||$$

and the **Cauchy-Schwarz Inequality**:

$$|\langle u, v\rangle| \leq ||u||||v||$$

where we have equality if and only if u and v are scalar multiples of each other. The Cauchy-Schwarz inequality is easily seen to be consistent with the definition of angle above. Finally, we have the **Pythagorean Theorem**, which states that

$$||u + v||^2 = ||u||^2 + ||v||^2 \text{ if and only if } u \text{ and } v \text{ are orthogonal}$$

Some examples of inner product spaces are as follows.

Example: $\mathbb{R}^n$ equipped with the **dot product**, given by

$$v \cdot w = \langle v, w\rangle = v_1 w_1 + v_2 w_2 + ... + w_n w_n$$

Example: $\ell^2(\mathbb{Z})$, the set of all two-sided real valued infinite sequences $..., a_{-2}, a_{-1}, a_0, a_1, a_2, ...$ with

$$\sum_{n\in\mathbb{Z}} |a_n|^2 < \infty$$

The natural inner product here for two sequences $\mathbf{a} = ..., a_{-2}, a_{-1}, a_0, a_1, a_2, ...$ and $\mathbf{b} = ..., b_{-2}, b_{-1}, b_0, b_1, b_2, ...$ is

$$\langle \mathbf{a}, \mathbf{b}\rangle = \sum_{n\in\mathbb{Z}} a_n b_n$$

This is an infinite dimensional inner product space.

Example: $L^2([0, 1])$ is the set of all real-valued **square integrable functions on** $[0, 1]$, meaning all functions f defined on $[0, 1]$ such that

$$\int_0^1 |f(x)|^2 dx < \infty$$

This is an inner product space over $\mathbb{R}$ with the inner product

$$\langle f, g\rangle = \int_0^1 f(x)g(x)dx$$

This is an infinite dimensional inner product space. What does the Cauchy-Schwarz inequality say in this case? Check that $f(x) = 1/2 - x$ and $g(x) = (1/2 - x)^2$ are orthogonal in this inner product space.

More on Orthogonality

We will restrict our attention to $\mathbb{R}^n$, or more generally, finite-dimensional inner product spaces. Recall that given an inner product space, we have a notion of perpendicularity or equivalently **orthogonality**, where $\langle v, w\rangle = 0$ if and only if v and w are orthogonal. Recall that we say that a vector v is a **unit vector** if $||v|| = 1$. We use these notions to define the following fundamental concepts.

Definition: Consider any finite set of vectors $v_1, v_2, ..., v_k$ in an inner product space V. We say that the vectors form an **orthogonal set** if $\langle v_i, v_j \rangle = 0$ whenever $i \neq j$ (so every distinct pair of vectors is orthogonal). If in addition, every vector in the set is a unit vector, we say that the vectors form an **orthonormal set**. If in addition, the vectors are a basis for V, the vectors are an **orthonormal basis** for V.

What is good about orthonormal sets? First, it is an important fact that *orthonormal sets are always linearly independent.* To see this, let $v_1, v_2, ..., v_k$ be an orthonormal set. Then, if

$$c_1v_1 + c_2v_2 + ... + c_kv_k = 0$$

then

$$\langle c_1v_1 + c_2v_2 + ... + c_kv_k, v_1 \rangle = \langle 0, v_1 \rangle = 0$$

which implies by bilinearity and orthonormality that $c_1 = 0$. A similar argument works for all c_i so that all $c_i = 0$. So orthonormal vectors are linearly independent.

Further, there is an additional benefit of having an orthonormal basis instead of just a basis. For a general basis, we have to solve systems of equations to get the coefficients for any linear combination. But for an orthonormal set, we can get the coefficients by taking inner products. For example, if $v_1, v_2, ..., v_n$ are an orthonormal basis for V, then given any vector w, if we want to find c_i such that

$$w = c_1v_1 + c_2v_2 + ... + c_nv_n$$

by taking the inner product with c_1 as above, we get that

$$c_1 = \langle w, v_1 \rangle$$

and more generally, that $c_i = \langle w, v_i \rangle$. This makes finding these coefficients much easier!

Given a subspace W of $\mathbb{R}^n$, we can define the **orthogonal complement** of W, denoted $W^\perp$, as the set of all vectors in V that are perpendicular to every vector in W.

Problem 1

Check that $W^\perp$ is a subspace of V. What is $W^\perp$ if $W = 0$? What is $W^\perp$ if $W = V$?

The only vector that is in both W and $W^\perp$ is the zero vector. (Why?) Using this fact, we can show that every vector in W is "linearly independent" from every vector in $W^\perp$. We notate this by saying $V = W \oplus W^\perp$, where this operation is called a direct sum. This means that every vector v in V can be written uniquely as a sum $v = w + u$ where $w \in W$ and $u \in W^\perp$. This also implies that

$$\dim(V) = \dim(W) + \dim(W^\perp)$$

Problem 2

Find the orthogonal complement of $W = \text{span}\{(1, 1, 2, 1), (1, 0, 0, -1)\}$ in $\mathbb{R}^4$.