Chapter 4

Inner product spaces

In this chapter we study a special class of normed spaces – those whose norms are induced by *inner products*. These spaces are well behaved in the sense that they share with $\mathbb{R}^2$ certain geometrically desirable properties. For instance, the standard norm on $\mathbb{R}^2$ obeys the parallelogram law:

$$\|\mathbf{x}-\mathbf{y}\|^2 + \|\mathbf{x}+\mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2), \tag{4.0.1}$$

where $\mathbf{x}$ and $\mathbf{y}$ are adjacent sides of the parallelogram, so that $\|\mathbf{x}-\mathbf{y}\|$ is the length of one diagonal and $\|\mathbf{x}+\mathbf{y}\|$ is the length of the other.

This law does not hold in an arbitrary normed space. Consider $C[0,1]$ with the sup-norm. For all $t \in [0,1]$, let $x(t) := t$, $y(t) := 1-t$. Then $x, y \in C[0,1]$ but $\|x-y\|^2 + \|x+y\|^2 = 2$ and $2(\|x\|^2 + \|y\|^2) = 4$.

But norms induced by an inner product do satisfy the parallelogram law.

Other desirable properties are restricted to a special class of inner product spaces: complete inner product spaces, called *Hilbert spaces*. For instance, let M be the x-axis in $\mathbb{R}^2$, and let $\mathbf{p}$ be the point on the y-axis where $y = 1$. In $\mathbb{R}^2$ with the standard norm, the origin is the unique point in M closest to $\mathbf{p}$, and $\|\mathbf{p}-\mathbf{0}\| = 1$. In $\mathbb{R}^2$ with the norm $\|\mathbf{x}\|_\infty = \max\{|x_1|, |x_2|\}$, this isn't true: *any* point on the x-axis in the interval $[-1,1]$ is distance 1 from $\mathbf{p}$. But Hilbert spaces behave in this regard like $\mathbb{R}^2$ with the standard norm.

Why is this desirable? Imagine that we want to approximate a function $x \in X$ and that $M \subset X$ consists of polynomials of degree at most n. Then the distance from x to M represents how well we can approximate the function by a polynomial in M. Clearly it is interesting to know whether there is a unique polynomial that provides the best possible approximation.

4.1 Definitions and examples

We denote by $\bar{t}$ the complex conjugate of a complex number t: if $t = a + bi$, where $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$, then $\bar{t} = a - bi$. Of course, $t = \bar{t}$ if and only if t is real.

Definition 4.1.1 (Inner product). Let X be a vector space over $\mathcal{K}$ (either $\mathbb{R}$ or $\mathbb{C}$). An *inner product* on X is a *function*

$$\langle\ ,\ \rangle : X \times X \to \mathcal{K}$$

that assigns to each pair $(x, y) \in X^2$ a number in $\mathcal{K}$, denoted $\langle x, y\rangle$, satisfying the following properties.

1. Conjugate symmetry: $\langle x, y\rangle = \overline{\langle y, x\rangle}$.
2. Linearity with respect to the first variable: for all $a, b \in \mathcal{K}$,

$$\langle ax_1 + bx_2,\ y\rangle = a\langle x_1, y\rangle + b\langle x_2, y\rangle.$$

3. Positivity: $\langle x, x\rangle \geq 0$; moreover, $\langle x, x\rangle = 0$ if and only if $x = \mathbf{0}$.

The pair $(X, \langle\ ,\ \rangle)$ is an *inner product space* over $\mathcal{K}$. If $\mathcal{K} = \mathbb{C}$, it is a complex inner product space; if $\mathcal{K} = \mathbb{R}$, it is a real inner product space. An inner product space is finite dimensional if the vector space X is finite dimensional. Otherwise, it is infinite dimensional. When it is clear what the inner product is, we may simply write "X".

Remarks 4.1.2.

1. Property 1 implies that $\langle x, x\rangle$ is real, so that property 3 makes sense; complex numbers that are not real are neither positive nor negative.
2. In some treatments, particularly by physicists, the inner product is linear with respect to the *second* variable:

$$\langle x, ay_1 + by_2\rangle = a\langle x, y_1\rangle + b\langle x, y_2\rangle. \quad \triangle$$

Just as the metric on a set was motivated by the "distance function" on $\mathbb{R}$, the inner product was motivated by the dot product. A common theme in mathematics is abstraction – isolating the essential properties of a concept in a concrete setting and using them to extend the concept to a more general setting. The hardest part is identifying the essential properties.

Example 4.1.3 (Dot product as inner product).

1. Define $\langle\ ,\ \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by

$$\langle \mathbf{x}, \mathbf{y}\rangle := x_1y_1 + \cdots + x_ny_n. \tag{4.1.1}$$

Then $\langle\ ,\ \rangle$, the dot product on $\mathbb{R}^n$, is an inner product on $\mathbb{R}^n$, and $(\mathbb{R}^n, \langle\ ,\ \rangle)$ is an inner product space over $\mathbb{R}$.

2. Define $\langle\ ,\ \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ by

$$\langle \mathbf{z}, \mathbf{w}\rangle := z_1\overline{w_1} + \cdots + z_n\overline{w_n}. \tag{4.1.2}$$

Then $(\mathbb{C}^n, \langle\ ,\ \rangle)$ is an inner product space over $\mathbb{C}$. Note the complex conjugate.

Whenever we consider $\mathbb{R}^n$ and $\mathbb{C}^n$ as inner product spaces, these are the inner products we mean. $\triangle$

Example 4.1.4 ($CL^2[a,b]$). Let $a < b$. Define $\langle\, , \rangle : C[a,b] \times C[a,b] \to \mathbb{C}$ by

$$\langle f, g \rangle := \int_a^b f(t)\overline{g}(t)\, dt. \tag{4.1.3}$$

Then $(C[a,b], \langle\, , \rangle)$ is an inner product space. We will denote this space $CL^2[a,b]$, the C to suggest continuity and L^2 because that is standard notation for the normed space of "square integrable" functions: functions f such that $\int_a^b |f(t)|^2 < \infty$. $\triangle$

Example 4.1.5 (l^2 as an inner product space). Recall that the set l^2 consists of sequences $(x_j) \subset \mathbb{C}$ such that $\sum_{j=1}^{\infty} |x_j|^2 < \infty$. By Hölder's inequality (Inequality 0.4),

$$\sum_{j=1}^{\infty} |x_j|\, |\overline{y_j}| < \infty \text{ for all } \mathbf{x} = (x_j), \mathbf{y} = (y_j) \in l^2. \tag{4.1.4}$$

We know from calculus that for a sequence of complex numbers, absolute convergence implies convergence, so we can define $\langle\, , \rangle : l^2 \times l^2 \to \mathbb{C}$ by

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{j=1}^{\infty} x_j\, \overline{y_j}. \tag{4.1.5}$$

Then $(l^2, \langle\, , \rangle)$ is an inner product space. Whenever we consider l^2 as an inner product space, this is the inner product we mean. $\triangle$

Example 4.1.6. Let Mat (n, m) denote the vector space of $n \times m$ matrices with real entries, and denote by tr the *trace* of a square matrix, i.e., the sum of the entries on the main diagonal. For example, $\operatorname{tr}\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} = a_1 + b_2$. Then if A and B are elements of Mat (n, m),

$$\langle A, B \rangle := \operatorname{tr}(AB^\top) \tag{4.1.6}$$

defines an inner product on Mat (n, m), as you are asked to show in Exercise 4.1.24.

Proposition 4.1.7 (Basic properties of inner products). Let $(X, \langle\, , \rangle)$ be an inner product space over $\mathcal{K}$, with x, y, y_1, $y_2 \in X$ and $a, b \in \mathcal{K}$. Then

1. $\langle \mathbf{0}, y \rangle = 0 = \langle x, \mathbf{0} \rangle$.
2. $\langle x,\, ay_1 + by_2 \rangle = \bar{a}\langle x, y_1 \rangle + \bar{b}\langle x, y_2 \rangle$.
3. $\langle ax, ax \rangle = |a|^2 \langle x, x \rangle$.
4. If $x_0, z_0 \in X$ and $\langle x_0, y \rangle = \langle z_0, y \rangle$ for all $y \in X$, then $x_0 = z_0$.

   In particular, $\langle x_0, y \rangle = 0$ for all $y \in X$ if and only if $x_0 = \mathbf{0}$.

Remark. Any function satisfying property 2 is said to be is *conjugate-linear* with respect to the second variable. △

PROOF. 1. $\langle \mathbf{0}, y\rangle = \langle 0\mathbf{0}, y\rangle = 0\langle \mathbf{0}, y\rangle = 0$ for all $y \in X$. Thus $\langle x, \mathbf{0}\rangle = \overline{\langle \mathbf{0}, x\rangle} = 0$ for all $x \in X$.

2. Just compute:

$$\langle x, ay_1 + by_2\rangle = \overline{\langle ay_1 + by_2, x\rangle} = \bar{a}\,\overline{\langle y_1, x\rangle} + \bar{b}\,\overline{\langle y_2, x\rangle} = \bar{a}\langle x, y_1\rangle + \bar{b}\langle x, y_2\rangle.$$

3. By part 2, $\langle ax, ax\rangle = \bar{a}\langle ax, x\rangle = \bar{a}a\langle x, x\rangle = |a|^2\langle x, x\rangle$.

4. If $\langle x_0, y\rangle = \langle z_0, y\rangle$ for all $y \in X$, then $\langle x_0 - z_0, y\rangle = 0$ for all $y \in X$. In particular,

$$\langle x_0 - z_0,\, x_0 - z_0\rangle = 0. \tag{4.1.7}$$

Thus, $x_0 - z_0 = \mathbf{0}$, hence, $x_0 = z_0$. The remaining assertion follows easily because

$$\langle x_0, y\rangle = 0 \text{ for all } y \in X \iff \langle x_0, y\rangle = \langle \mathbf{0}, y\rangle \text{ for all } y \in X \iff x_0 = \mathbf{0}. \qquad \square$$

Complex inner product spaces are generally easier to deal with than real inner product spaces. This is illustrated by the next example.

Example 4.1.8. Let $(X, \langle\ ,\ \rangle)$ be a complex inner product space, and let $Q : X \to X$ be any linear map such that $\langle Qv, v\rangle = 0$ for all $v \in X$. Let us see that $Q = \mathbf{0}$, the zero map. For all $x, y \in X$ and all $\alpha \in \mathbb{C}$, we have

$$\begin{aligned} 0 &= \langle Q(\alpha x + y), \alpha x + y\rangle = \langle Q(\alpha x) + Qy,\ \alpha x + y\rangle \\ &= \underbrace{\langle Q(\alpha x), \alpha x\rangle}_{0} + \langle Q(\alpha x), y\rangle + \langle Qy, \alpha x\rangle + \underbrace{\langle Qy, y\rangle}_{0} \\ &= \alpha\langle Qx, y\rangle + \bar{\alpha}\langle Qy, x\rangle. \end{aligned} \tag{4.1.8}$$

Put first $\alpha = 1$ and then $\alpha = i$ in the equation $0 = \alpha\langle Qx, y\rangle + \bar{\alpha}\langle Qy, x\rangle$ to deduce that $\langle Qy, x\rangle = 0$ for all $x, y \in X$. Hence, by Proposition 4.1.7, $Qy = \mathbf{0}$ for all $y \in X$, so $Q = \mathbf{0}$. Note that this is not the case when $(X, \langle\ ,\ \rangle)$ is a real inner product space. For instance, let $Q : \mathbb{R}^2 \to \mathbb{R}^2$ rotate each $\mathbf{x} \in \mathbb{R}^2$ by 90 degrees. △

Norms induced by inner products

Note that the Euclidean norm of a vector $\mathbf{x} \in \mathbb{R}^n$ is the square root of the standard dot product of $\mathbf{x}$ with itself:

$$\|\mathbf{x}\| = \sqrt{x_1^2 + \cdots x_n^2} = \sqrt{\langle \mathbf{x}, \mathbf{x}\rangle}, \tag{4.1.9}$$