Axler: Inner Product Spaces

James Pagan

January 2023

Contents

1 Inner Products and Norms

1.1 Inner Products

An **inner product** over a complex (or real) vector space V is a map $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{C}$ that satisfies the following properties for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{V}$ and $\lambda \in \mathbb{C}$:

1. **Conjugate Symmetry**: $\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$
2. **Positive-Definiteness**: $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$, with equality if and only if $\mathbf{v} = \mathbf{0}$.
3. **Additivity in First Argument**: $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$.
4. **Homogenity in First Argument**: $\langle \lambda \mathbf{v}, \mathbf{w} \rangle = \lambda \langle \mathbf{v}, \mathbf{w} \rangle$.

As $z = \overline{z}$ if and only if z is real, (1) implies that $\langle \mathbf{v}, \mathbf{v} \rangle \in \mathbb{R}$; hence, (3) is a valid condition. An **inner product space** is a vector space V over $\mathbb{R}$ or $\mathbb{C}$. We will exclusively prove theorems about complex vector spaces; proofs for inner product spaces over $\mathbb{R}$ are identical.

Theorem 1. *Suppose V is an inner product space. Then the following five properties hold: for each $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $\lambda \in \mathbb{C}$:*

1. *The function $f(\mathbf{v}) = \langle \mathbf{v}, \mathbf{w} \rangle$ is a linear map from V to $\mathbb{C}$.*
2. $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$.
3. $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$.
4. $\langle \mathbf{v}, \lambda \mathbf{w} \rangle = \overline{\lambda} \langle \mathbf{v}, \mathbf{w} \rangle$

Proof. Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be arbitrary vectors of V and let λ be an arbitrary scalar of $\mathbb{C}$. For (1), note that

$$f(\mathbf{u} + \mathbf{v}) = \langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle = f(\mathbf{u}) + f(\mathbf{v})j$$

and see that

$$f(\lambda \mathbf{v}) = \langle \lambda \mathbf{v}, \mathbf{w} \rangle = \lambda \langle \mathbf{v}, \mathbf{w} \rangle = \lambda f(\mathbf{v});$$

thus f is linear. For (2), observe that

$$\langle \mathbf{0}, \mathbf{v} \rangle = \langle 0\mathbf{v}, \mathbf{v} \rangle = 0 \langle \mathbf{v}, \mathbf{v} \rangle = 0.$$

Similarly, $\langle \mathbf{v}, \mathbf{0} \rangle = \overline{\langle \mathbf{0}, \mathbf{v} \rangle} = \overline{0} = 0$. For (3), notice that

$$\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \overline{\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle} = \overline{\langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle} = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle.$$

Finally, we have that $\langle \mathbf{v}, \lambda \mathbf{w} \rangle = \overline{\langle \lambda \mathbf{w}, \mathbf{v} \rangle} = \overline{\lambda \langle \mathbf{w}, \mathbf{v} \rangle} = \overline{\lambda} \langle \mathbf{v}, \mathbf{w} \rangle$, yielding (4). ☐

1.2 Norms

The **norm** of a vector $\mathbf{v}$ in an inner product space V is defined as

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

We may also define normed vector spaces, which utilize a norm without an inner product—but this lies beyond the scope of this document.

Theorem 2. *Suppose V is an inner product space. Then the following properties hold for all $\mathbf{v} \in V$ and $\lambda \in \mathbb{C}$:*

- $\|\mathbf{v}\| = 0$ *if and only if* $\mathbf{v} = \mathbf{0}$.
- $\|\lambda \mathbf{v}\| = |\lambda| \|\mathbf{v}\|$.

Proof. (1) follows from the fact that

$$\begin{aligned} \|\mathbf{v}\| = 0 &\iff \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = 0 \\ &\iff \langle \mathbf{v}, \mathbf{v} \rangle = 0 \\ &\iff \mathbf{v} = 0. \end{aligned}$$

For (2), see that

$$\|\lambda \mathbf{v}\| = \sqrt{\langle \lambda \mathbf{v}, \lambda \mathbf{v} \rangle} = \sqrt{\lambda \bar{\lambda} \langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{|\lambda|^2 \langle \mathbf{v}, \mathbf{v} \rangle} = |\lambda| \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = |\lambda| \|\mathbf{v}\|,$$

as desired. $\square$

Two vectors $\mathbf{v}$ and $\mathbf{w}$ in an inner product space V are **orthogonal** if $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. Clearly, every vector is orthogonal to $\mathbf{0}$ — and the only vector orthogonal to itself is also $\mathbf{0}$.

Theorem 3 (Pythagorean Theorem)**.** *Suppose V is an inner product space. Then if $\mathbf{v}, \mathbf{w} \in V$ are orthogonal, then*

$$\|\mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2.$$

Proof. Suppose $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. Then

$$\begin{aligned} \|\mathbf{v} + \mathbf{w}\|^2 &= \langle \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w} \rangle \\ &= \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle \\ &= \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle \\ &= \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2. \end{aligned}$$

This is like the Pythagorean Theorem for vectors. $\square$

Theorem 4. *Suppose* $\mathbf{v} \neq \mathbf{0}$ *and* $\mathbf{w}$ *are vectors in an inner product space. Then setting* $c = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{w}\|^2}$ *and* $\mathbf{u} = \mathbf{v} - \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{w}\|^2}\mathbf{w}$ *yields*

$$\mathbf{v} = c\mathbf{w} + \mathbf{u} \qquad \text{and} \qquad \langle \mathbf{u}, \mathbf{w} \rangle = 0.$$

Proof. The result is a mere computation. Clearly $\mathbf{v} = c\mathbf{w} + \mathbf{u}$; as for the orthogonality,

$$\begin{aligned}
\langle \mathbf{u}, \mathbf{w} \rangle &= \left\langle \mathbf{v} - \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{w}\|^2}\mathbf{w}, \mathbf{w} \right\rangle \\
&= \langle \mathbf{v}, \mathbf{w} \rangle - \left\langle \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{w}\|^2}\mathbf{w}, \mathbf{w} \right\rangle \\
&= \langle \mathbf{v}, \mathbf{w} \rangle - \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{w}\|^2} \langle \mathbf{w}, \mathbf{w} \rangle \\
&= \langle \mathbf{v}, \mathbf{w} \rangle - \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{w}\|^2} \|\mathbf{w}\|^2 \\
&= \langle \mathbf{v}, \mathbf{w} \rangle - \langle \mathbf{v}, \mathbf{w} \rangle \\
&= 0,
\end{aligned}$$

as required. The vector $c\mathbf{w}$ is often denoted $\mathrm{Proj}_{\mathbf{v}}(\mathbf{w})$. □

The following inequality is the most important in all of mathematics.

Theorem 5 (Cauchy-Schwarz Inequality)**.** *Suppose* V *is an inner product space. Then for all* $\mathbf{v}, \mathbf{w} \in V$,

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \|\mathbf{v}\| \|\mathbf{w}\|,$$

with equality if and if one of $\mathbf{v}, \mathbf{w}$ *is a scalar multiple of the other.*

Proof. Enabled by Theorem 4, we consider the orthogonal decomposition below, defining $\mathbf{u}$ in the process:

$$\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{w}\|^2}\mathbf{w} + \mathbf{u}.$$

For simplicity, let $\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{w}\|^2} = c$. Now $\langle c\mathbf{w}, \mathbf{u} \rangle = c \langle \mathbf{w}, \mathbf{u} \rangle = 0$, so by the Pythagorean

Theorem,

$$\begin{aligned}
\|\mathbf{v}\|^2\|\mathbf{w}\|^2 &= \|c\mathbf{w}+\mathbf{u}\|^2\|\mathbf{w}\|^2 \\
&= (|c|^2\|\mathbf{w}\|^2 + \|\mathbf{u}\|^2)\mathbf{v}^2 \\
&= \left(\frac{|\langle \mathbf{v},\mathbf{w}\rangle|^2}{\|\mathbf{w}\|^4}\|\mathbf{w}\|^2 + \|\mathbf{u}\|^2\right)\|\mathbf{w}\|^2 \\
&= \frac{|\langle \mathbf{v},\mathbf{w}\rangle|^2}{\|\mathbf{w}\|^2}\|\mathbf{w}\|^2 + \|\mathbf{u}\|^2\|\mathbf{w}\|^2 \\
&= |\langle \mathbf{v},\mathbf{w}\rangle|^2 + \|\mathbf{u}\|^2\|\mathbf{w}\|^2 \\
&\geq |\langle \mathbf{v},\mathbf{w}\rangle|^2.
\end{aligned}$$

We achieve equality at the last step if $\mathbf{w} = \mathbf{0}$ or if $\mathbf{u} = \mathbf{0}$; that is, if there exists $c \in \mathbb{C}$ such that $\mathbf{v} = c\mathbf{w}$. In either case, $\mathbf{v}$ and $\mathbf{w}$ are scalar multiples of each other. This proves the Cauchy-Schwarz Inequality. $\square$

Theorem 6 (Triangle Inequality)**.** *Suppose V is an inner product space and $\mathbf{v}, \mathbf{w} \in V$. Then*

$$\|\mathbf{v}\| + \|\mathbf{w}\| \geq \|\mathbf{v}+\mathbf{w}\|,$$

with equality if and only if one of $\mathbf{v}, \mathbf{w}$ is a nonnegative real multiple of the other.

Proof. We have that

$$\begin{aligned}
\|\mathbf{v}\| + \|\mathbf{w}\| &= \sqrt{(\|\mathbf{v}\| + \|\mathbf{w}\|)^2} \\
&= \sqrt{\|\mathbf{v}\|^2 + 2\|\mathbf{v}\|\|\mathbf{w}\| + \|\mathbf{w}\|^2} \\
&\geq \sqrt{\|\mathbf{v}\|^2 + 2|\langle \mathbf{v},\mathbf{w}\rangle| + \|\mathbf{w}\|^2} \\
&\geq \sqrt{\|\mathbf{v}\|^2 + 2\operatorname{Re}\langle \mathbf{v},\mathbf{w}\rangle + \|\mathbf{w}\|^2} \\
&= \sqrt{\|\mathbf{v}\|^2 + \langle \mathbf{v},\mathbf{w}\rangle + \overline{\langle \mathbf{v},\mathbf{w}\rangle} + \|\mathbf{w}\|^2} \\
&= \sqrt{\langle \mathbf{v},\mathbf{v}\rangle + \langle \mathbf{v},\mathbf{w}\rangle + \langle \mathbf{w},\mathbf{v}\rangle + \langle \mathbf{w},\mathbf{w}\rangle} \\
&= \sqrt{\langle \mathbf{v}+\mathbf{w},\mathbf{v}+\mathbf{w}\rangle} \\
&= \sqrt{\|\mathbf{v}+\mathbf{w}\|^2} \\
&= \|\mathbf{v}+\mathbf{w}\|^2,
\end{aligned}$$

as required. Equality holds in the first inequality if and only if one of $\mathbf{v}, \mathbf{w}$ is a scalar multiple of the other; if this is the case, then $\langle \mathbf{v},\mathbf{w}\rangle$ is a scalar multiple of $\langle \mathbf{v},\mathbf{v}\rangle$.

The second inequality holds if this scalar multiple is positive — proving the Triangle Inequality. $\square$

Theorem 7 (Parallelogram Equality)**.** *Suppose V is an inner product space and $\mathbf{v}, \mathbf{w} \in V$. Then*

$$\|\mathbf{v} + \mathbf{w}\|^2 + \|\mathbf{v} - \mathbf{w}\|^2 = 2\left(\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2\right).$$

Proof. We have that

$$\begin{aligned}
\|\mathbf{v} + \mathbf{w}\|^2 + \|\mathbf{v} - \mathbf{w}\|^2 &= \langle \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w} \rangle + \langle \mathbf{v} - \mathbf{w}, \mathbf{v} - \mathbf{w} \rangle \\
&= \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle \\
&\quad + \langle \mathbf{v}, \mathbf{v} \rangle - \langle \mathbf{v}, \mathbf{w} \rangle - \langle \mathbf{w}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle \\
&= 2\langle \mathbf{v}, \mathbf{v} \rangle + 2\langle \mathbf{w}, \mathbf{w} \rangle \\
&= 2\|\mathbf{v}\|^2 + 2\|\mathbf{w}\|^2,
\end{aligned}$$

as required. $\square$

2 Orthonormal Bases

A list of vectors is called **orthonormal** if each vector in the list has norm 1 and is orthogonal to all the other vectors in the list. In other words, a list of vectors $\mathbf{e}_1, \ldots, \mathbf{e}_m \in V$ is orthonormal if

$$\langle \mathbf{e}_j, \mathbf{e}_k \rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k \end{cases}$$

for all $j, k \in \{1, \ldots, m\}$.

Theorem 8. *Suppose $\mathbf{e}_1, \ldots, \mathbf{e}_n \in V$ is an orthonormal list. Then for all $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$,*

$$\|\lambda_1 \mathbf{e}_1 + \cdots + \lambda_m \mathbf{e}_m\|^2 = |\lambda_1|^2 + \cdots + |\lambda_m|^2.$$

Proof. Realize that $\langle \lambda_j \mathbf{e}_j, \lambda_k \mathbf{e}_k \rangle = \lambda_j \overline{\lambda_k} \langle \mathbf{e}_j, \mathbf{e}_k \rangle = 0$ for all $j, k \in \{1, \ldots, m\}$ and $j \neq k$. Then by the Pythagorean Theorem,

$$\|\lambda_1 \mathbf{e}_1 + \cdots + \lambda_m \mathbf{e}_m\|^2 = \|\lambda_1 \mathbf{e}_1\|^2 + \cdots + \|\lambda_m \mathbf{e}_m\|^2 = |\lambda_1|^2 + \cdots + |\lambda_m|^2,$$

as desired. $\square$

Theorem 9. *Every orthonormal list of vectors is linearly independent.*

Proof. Let $\mathbf{e}_1, \ldots, \mathbf{e}_m \in V$ be an orthonormal list of vectors. Suppose for contradiction that there exist $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$, not all zero, such that

$$\lambda_1 \mathbf{e}_1 + \cdots + \lambda_m \mathbf{e}_m = \mathbf{0}.$$

Then by Theorem 8,

$$0 = \|\lambda_1 \mathbf{e}_1 + \cdots + \lambda_m \mathbf{e}_m\|^2 = |\lambda_1|^2 + \cdots + |\lambda_m|^2.$$

We conclude that all the $\lambda_1, \ldots, \lambda_m$ are zero, which yields the desired contradiction. □

Theorem 10. *Suppose $\mathbf{e}_1, \ldots, \mathbf{e}_m \in V$ is an orthonormal list of vectors. Then for all $\mathbf{v} \in V$,*

$$|\langle \mathbf{v}, \mathbf{e}_1 \rangle|^2 + \cdots + |\langle \mathbf{v}, \mathbf{e}_m \rangle|^2 \leq \|\mathbf{v}\|^2.$$

Proof. I am not sure, but I would like to think of a proof myself. □

An **orthonormal basis** of V is an orthonormal list of vectors in V that is a basis of V. If V is finite dimensional, then any orthonormal list of length $\dim V$ is an orthonrmal basis.

Each $\langle \mathbf{v}, \mathbf{e}_i \rangle$ for $i \in \{1, \ldots, n\}$ equals the i-th coordinate of $\mathbf{v}$ as written as a lienar combination of $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — an idea expanded upon in the following theorem:

Theorem 11. *Suppose $\mathbf{e}_1, \ldots, \mathbf{e}_n \in V$ is an orthonormal basis of V and $\mathbf{v}, \mathbf{w} \in V$. Then the following three identities hold:*

1. $\mathbf{v} = \langle \mathbf{v}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \cdots + \langle \mathbf{v}, \mathbf{e}_n \rangle \mathbf{e}_n$.
2. $\|\mathbf{v}\|^2 = |\langle \mathbf{v}, \mathbf{e}_1 \rangle|^2 + \cdots + |\langle \mathbf{v}, \mathbf{e}_n \rangle|^2$.
3. $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{e}_1 \rangle \overline{\langle \mathbf{w}, \mathbf{e_1} \rangle} + \cdots + \langle \mathbf{v}, \mathbf{e}_n \rangle \overline{\langle \mathbf{w}, \mathbf{e_n} \rangle}$

Proof. Define $v_1, \ldots, v_n \in \mathbb{C}$ and $w_1, \ldots, w_n \in \mathbb{C}$ such that

$$v_1 \mathbf{e}_1 + \cdots + v_n \mathbf{e}_n = \mathbf{v} \quad \text{and} \quad w_1 \mathbf{e}_1 + \cdots + w_n \mathbf{e}_n = \mathbf{w}.$$

For (1), realize that for all $i \in \{1, \ldots, n\}$,

$$\langle \mathbf{v}, \mathbf{e}_i \rangle = \langle v_1 \mathbf{e}_1 + \cdots + v_n \mathbf{e}_n, \mathbf{e}_i \rangle + v_1 \langle \mathbf{e}_1, \mathbf{e}_i \rangle + \cdots + v_n \langle \mathbf{e}_n, \mathbf{e}_i \rangle = v_i.$$

Therefore,

$$\langle \mathbf{v}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \cdots + \langle \mathbf{v}, \mathbf{e}_n \rangle \mathbf{e}_n = v_1 \mathbf{e}_1 + \cdots v_n \mathbf{e}_n = \mathbf{v}.$$

(2) follows immmediately from Theorem 8. For (3), we may use the formula of (1) to derive that

$$\begin{aligned}
\langle \mathbf{v}, \mathbf{w} \rangle &= \langle v_1 \mathbf{e}_1 + \cdots + v_n \mathbf{e_n}, \mathbf{w} \rangle \\
&= v_1 \langle \mathbf{e}_1, \mathbf{w} \rangle + \cdots + v_n \langle \mathbf{e}_n, \mathbf{w} \rangle \\
&= v_1 \overline{\langle \mathbf{w}, \mathbf{e}_1 \rangle} + \cdots + v_n \overline{\langle \mathbf{w}, \mathbf{e}_n \rangle} \\
&= \langle \mathbf{v}, \mathbf{e}_1 \rangle \overline{\langle \mathbf{w}, \mathbf{e_1} \rangle} + \cdots + \langle \mathbf{v}, \mathbf{e}_n \rangle \overline{\langle \mathbf{w}, \mathbf{e_n} \rangle},
\end{aligned}$$

as required. $\square$