

11.1 Inner Products and Norms

Nine chapters, and we have not measured anything. We can say that a list is independent, that a subspace has dimension four, that an operator is diagonalizable. We cannot yet say that a vector has length 3, or that two vectors meet at a right angle. Geometry has appeared only in pictures, never in the theory. This section puts it into the theory, by abstracting the one operation in $\mathbb{R}^n$ that already computes lengths and angles: the dot product.

11.1.1 The field for this chapter

Throughout Chapter 11 the field F is $\mathbb{R}$ or $\mathbb{C}$. This is not laziness; two features of these fields are used on every page, and no other field has both.

The first is **positivity**. The point of the whole construction is that $\langle v, v \rangle$ should be the square of a length, so it must be a number we can compare with 0. Over $\mathbb{F}_p$ there is no ordering compatible with the arithmetic at all: in $\mathbb{F}_5$, adding 1 to itself five times returns to 0, so no nonzero element can be called “positive” in any useful sense. Over $\mathbb{Q}$ there is an ordering, but lengths would not exist, since $\sqrt{2} \notin \mathbb{Q}$. Over $\mathbb{R}$ every sum of squares is ≥ 0 and every nonnegative number has a square root, and that is exactly what we need.

The second is **conjugation**. Over $\mathbb{C}$ the naive sum of squares fails immediately: $1^2 + i^2 = 0$, so “ $z_1^2 + z_2^2$ ” is not a length for $(1, i) \neq \mathbf{0}$. The repair is to use $z\bar{z} = |z|^2$ instead of z^2 (Theorem 0.6.2 (c)), and this needs the complex conjugate. We write $\bar{a}$ for the conjugate of $a \in F$, with the convention that $\bar{a} = a$ when $F = \mathbb{R}$; every formula below then reads correctly over both fields at once.

11.1.2 From the dot product to the axioms

In $\mathbb{R}^n$ the dot product is

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_n y_n.$$

It computes lengths, because $\mathbf{x} \cdot \mathbf{x} = x_1^2 + \cdots + x_n^2$ is the square of the Euclidean length by Pythagoras. It computes angles too. For nonzero $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ at an angle θ , the law of cosines applied to the triangle with sides $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{x} - \mathbf{y}$ says

$$\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta,$$

while expanding the left side entry by entry gives $\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2(\mathbf{x} \cdot \mathbf{y})$. Comparing the two, $\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\|\cos\theta$. So one bilinear expression carries both length and angle, and perpendicularity is the single equation $\mathbf{x} \cdot \mathbf{y} = 0$.

Which properties of the dot product did that argument use? Only three: it is linear in each slot separately, it is symmetric, and $\mathbf{x} \cdot \mathbf{x} > 0$ for $\mathbf{x} \neq \mathbf{0}$. Those three we can demand of an abstract space. Over $\mathbb{C}$, as we saw, symmetry has to be weakened to conjugate symmetry, and the price is small: a conjugate bar in the second slot.

An inner product is a way of multiplying two vectors into a scalar so that a vector paired with itself has positive “length squared”.

Definition 11.1.1. Inner Product

Let F be $\mathbb{R}$ or $\mathbb{C}$, and let V be a vector space over F . An **inner product** on V is a map

$$\langle \cdot, \cdot \rangle: V \times V \rightarrow F$$

such that for all $u, v, w \in V$ and all $a, b \in F$:

(IP1) $\langle au + bv, w \rangle = a\langle u, w \rangle + b\langle v, w \rangle$ (**linear in the first slot**);

(IP2) $\langle v, u \rangle = \overline{\langle u, v \rangle}$ (**conjugate symmetry**);

(IP3) $\langle v, v \rangle > 0$ for **every nonzero** $v \in V$ (**positive definiteness**).

A vector space equipped with an inner product is an **inner product space**. When $F = \mathbb{R}$ it is a **real** inner product space, when $F = \mathbb{C}$ a **complex** one.

In words, clause by clause. (IP1) says that if we freeze the second slot, what is left is a linear functional on V : the inner product distributes over sums and lets scalars out of the **first** argument untouched. (IP2) says that swapping the two arguments conjugates the answer; over $\mathbb{R}$ it is plain symmetry, $\langle v, u \rangle = \langle u, v \rangle$. (IP3) says that a nonzero vector pairs with itself to give a **strictly** positive number. For the inequality in (IP3) to mean anything, $\langle v, v \rangle$ must be a **real** number; it is, and the reason is (IP2), as the next result shows. Note also what (IP3) does **not** say: it says nothing about $\langle u, v \rangle$ for $u \neq v$, which may be negative, or not even real.

The first consequences are small, and we will use all of them constantly.

Theorem 11.1.1. First Properties of an Inner Product

Let V be an inner product space over F . For all $u, v, w \in V$ and $a, b \in F$:

- (a) $\langle v, v \rangle$ is a **real** number, and $\langle v, v \rangle \geq 0$ with equality if and only if $v = 0$;
- (b) $\langle u, av + bw \rangle = \bar{a}\langle u, v \rangle + \bar{b}\langle u, w \rangle$ (**conjugate-linear in the second slot**);
- (c) $\langle 0, v \rangle = \langle v, 0 \rangle = 0$;
- (d) if $\langle u, w \rangle = \langle v, w \rangle$ for **every** $w \in V$, then $u = v$.

Idea. Each part is one application of an axiom. For (a), a number equal to its own conjugate is real. For (b), swap slots with (IP2), use (IP1), and swap back. For (d), the only vector orthogonal to everything is the zero vector, and we find that out by pairing the difference **with itself**: that is the move to remember.

Proof.

- (a) By (IP2) with $u = v$, $\langle v, v \rangle = \overline{\langle v, v \rangle}$, and a complex number equal to its own conjugate has zero imaginary part, so $\langle v, v \rangle \in \mathbb{R}$. If $v \neq 0$, then $\langle v, v \rangle > 0$ by (IP3). If $v = 0$, then $\langle 0, 0 \rangle = \langle 0 \cdot 0, 0 \rangle = 0\langle 0, 0 \rangle = 0$ by (IP1). Hence $\langle v, v \rangle \geq 0$ always, with equality exactly when $v = 0$.

- (b) Using (IP2), then (IP1), then (IP2) again,

$$\begin{aligned} \langle u, av + bw \rangle &= \overline{\langle av + bw, u \rangle} \\ &= \overline{a\langle v, u \rangle + b\langle w, u \rangle} \\ &= \bar{a}\overline{\langle v, u \rangle} + \bar{b}\overline{\langle w, u \rangle} \\ &= \bar{a}\langle u, v \rangle + \bar{b}\langle u, w \rangle, \end{aligned}$$

where the third equality uses Theorem 0.6.2 (a) and (b).

- (c) $\langle 0, v \rangle = \langle 0 \cdot 0, v \rangle = 0\langle 0, v \rangle = 0$ by (IP1), and $\langle v, 0 \rangle = \overline{\langle 0, v \rangle} = 0$ by (IP2).

- (d) Put $z = u - v$. For every w , (IP1) gives $\langle z, w \rangle = \langle u, w \rangle - \langle v, w \rangle = 0$. Take $w = z$: then $\langle z, z \rangle = 0$, so $z = 0$ by (a), that is, $u = v$. ■

Part (d) is the first appearance of the move that runs through the whole chapter: **to show a vector is zero, pair it with itself**. There is exactly one vector of length zero, so an identity of the form $\langle z, z \rangle = 0$ is a proof that $z = 0$.

Part (b) is where the conjugate becomes visible in practice. An inner product over $\mathbb{C}$ is **not** bilinear: scalars leave the second slot wearing a bar. A map with properties (IP1) and (b) is called **sesquilinear**, from the Latin for “one and a half times linear”.

11.1.3 Examples

The dot product on $\mathbb{R}^n$. Define $\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \cdots + x_n y_n$. (IP1) holds because each term is linear in x_i ; (IP2) is the commutativity of multiplication in $\mathbb{R}$, and conjugation is the identity; (IP3) holds because $\langle \mathbf{x}, \mathbf{x} \rangle = x_1^2 + \cdots + x_n^2$ is a sum of squares of real numbers, which is 0 only when every $x_i = 0$. This is the **standard** inner product on $\mathbb{R}^n$, and it is the model for everything else.

The standard inner product on $\mathbb{C}^n$. Define

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 \overline{y_1} + \cdots + x_n \overline{y_n}.$$

(IP1) is clear. For (IP2), $\overline{\langle \mathbf{x}, \mathbf{y} \rangle} = \sum \overline{x_i y_i} = \sum \overline{x_i} \overline{y_i} = \langle \mathbf{y}, \mathbf{x} \rangle$ by Theorem 0.6.2 (a), (b). For (IP3), $\langle \mathbf{x}, \mathbf{x} \rangle = \sum x_i \overline{x_i} = \sum |x_i|^2$ by Theorem 0.6.2 (c), a sum of nonnegative **real** numbers, zero only when every $x_i = 0$. The conjugates are doing exactly one job: making the diagonal values real and nonnegative.

Weighted inner products. Fix real numbers $w_1, \dots, w_n > 0$ and put $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_i w_i x_i \overline{y_i}$ on F^n . The same three checks go through, with $\langle \mathbf{x}, \mathbf{x} \rangle = \sum w_i |x_i|^2$; the weights must be **strictly** positive, or (IP3) fails for a standard basis vector. In the degenerate case $n = 1$ and $w_1 = 1$, this is $\langle x, y \rangle = x \overline{y}$ on F itself, the smallest nonzero inner product space there is. It matters because it is where the axioms are just the arithmetic of $|\cdot|$.

Continuous functions. Let $V = C[0, 1]$ be the space of continuous functions $f: [0, 1] \rightarrow F$, and set

$$\langle f, g \rangle = \int_0^1 f(t) \overline{g(t)} dt.$$

(IP1) is linearity of the integral, and (IP2) follows because conjugation commutes with integration. For (IP3), $\langle f, f \rangle = \int_0^1 |f(t)|^2 dt \geq 0$, and we must rule out equality for $f \neq 0$. Here is the one fact from analysis we assume: if f is continuous and $f(c) \neq 0$ for some c , then $|f| \geq |f(c)|/2$ on some interval of positive length δ around c , so $\int_0^1 |f|^2 \geq \delta |f(c)|^2/4 > 0$. Continuity is not decoration; the nonexample below shows what happens without it.

Matrices: the Frobenius inner product. First a piece of notation we will need all chapter.

Definition 11.1.2. Conjugate Transpose

For $\mathbf{A} \in M_{m \times n}(F)$, the **conjugate transpose** $\mathbf{A}^* \in M_{n \times m}(F)$ is the matrix with entries $(\mathbf{A}^*)_{ij} = \overline{a_{ji}}$. A square matrix with $\mathbf{A}^* = \mathbf{A}$ is **Hermitian**.

Over $\mathbb{R}$, $\mathbf{A}^* = \mathbf{A}^\top$ and Hermitian means symmetric (Definition 0.9.11). Now on $V = M_{m \times n}(F)$ define

$$\langle \mathbf{A}, \mathbf{B} \rangle = \text{tr}(\mathbf{B}^* \mathbf{A}).$$

Writing out the trace, $\text{tr}(\mathbf{B}^* \mathbf{A}) = \sum_j \sum_i \overline{b_{ij}} a_{ij}$, so this is the standard inner product of F^{mn} after listing the entries in a column. In particular (IP1)–(IP3) hold, with $\langle \mathbf{A}, \mathbf{A} \rangle = \sum_{i,j} |a_{ij}|^2$. This is the **Frobenius inner product**; for $n = 1$ it is the standard inner product on F^m .

Polynomials sampled at nodes. Fix $n + 1$ **distinct** scalars $c_0, \dots, c_n \in F$ and define, on $V = F[x]_{\leq n}$,

$$\langle p, q \rangle = \sum_{i=0}^n p(c_i) \overline{q(c_i)}.$$

(IP1) and (IP2) hold because evaluation at c_i is linear (Theorem 0.8.5) and the standard inner product on F^{n+1} has them; indeed $\langle p, q \rangle$ is the standard inner product of the sample vectors $(p(c_0), \dots, p(c_n))$ and $(q(c_0), \dots, q(c_n))$. Positivity is the interesting clause. If $\langle p, p \rangle = \sum |p(c_i)|^2 = 0$, then p vanishes at all $n + 1$ distinct nodes. A nonzero polynomial of degree at most n has at most n distinct roots (Corollary 6.1.4), so $p = 0$. The **number** of nodes is what makes this work, and the nonexample below removes one.

11.1.4 Nonexamples, by minimal change

Change a sign: the hyperbolic form on $\mathbb{R}^2$. Let $\beta(\mathbf{x}, \mathbf{y}) = x_1 y_1 - x_2 y_2$. Bilinearity survives, since each term is still linear in each slot, and so does symmetry. What fails is (IP3), and the witness is $\mathbf{e}_2$: $\beta(\mathbf{e}_2, \mathbf{e}_2) = -1 < 0$. The vector $(1, 1)$ is worse still: it is nonzero, yet $\beta((1, 1), (1, 1)) = 0$, so it would be a nonzero vector of length 0. Forms like β are useful (they are the subject of the chapter on bilinear forms), but they are not inner products.

Drop the conjugate on $\mathbb{C}^2$. Let $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^\top \mathbf{v} = u_1 v_1 + u_2 v_2$ on $\mathbb{C}^2$. This map is linear in **both** slots and symmetric, so it satisfies (IP1) and looks tidier than the standard product. Take $\mathbf{u} = (1, i)$. Then

$$\langle \mathbf{u}, \mathbf{u} \rangle = 1^2 + i^2 = 1 - 1 = 0,$$

with $\mathbf{u} \neq \mathbf{0}$. Clause (IP3) fails. This one example is the whole reason for the conjugate: over $\mathbb{C}$, sums of squares vanish on nonzero vectors, sums of squared **moduli** do not.

Enlarge the space: drop continuity. Keep the formula $\langle f, g \rangle = \int_0^1 f g$, but let V be all Riemann-integrable functions $[0, 1] \rightarrow \mathbb{R}$. Linearity and symmetry are untouched. Let f be 1 at $t = \frac{1}{2}$ and 0 elsewhere. It is integrable, nonzero as a function, and $\int_0^1 f^2 = 0$. Positivity fails, so $\langle \cdot, \cdot \rangle$ is only **positive semidefinite** here. There are two standard repairs: restrict to continuous functions, as above, or regard two functions as equal when they differ on a negligible set. We take the first, and every function space in this chapter will be a space of continuous functions.

Remove a node. On $\mathbb{R}[x]_{\leq 2}$, put $\langle p, q \rangle = p(0)q(0) + p(1)q(1)$: the node example with one node too few. Bilinearity and symmetry hold. But $p = x^2 - x$ is a nonzero polynomial of degree 2 with $p(0) = p(1) = 0$, so $\langle p, p \rangle = 0$. The clause that fails is (IP3), and it fails because two nodes cannot pin down a quadratic.

Why this definition. Each clause earns its place. Without (IP3) there can be nonzero vectors v with $\langle v, v \rangle = 0$, and the proof of [Theorem 11.1.1](#) (d) breaks down, because pairing a vector with itself no longer detects zero. (The form $x_1 y_1 - x_2 y_2$ below shows that the conclusion of (d) may survive anyway; what is lost is the argument for it, and with it every later proof that runs through $\langle v, v \rangle = 0 \Rightarrow v = 0$.) Without (IP2) the diagonal values need not be real, and “ $\langle v, v \rangle > 0$ ” is meaningless. Linearity in the first slot rather than the second is a convention, but

once chosen it must be kept: with (IP2) in place, linearity in one slot forces conjugate-linearity in the other, so a map cannot be linear in both unless $F = \mathbb{R}$.

Warning. The conjugate goes in the second slot here, and books differ. With our convention, $\langle au, v \rangle = a\langle u, v \rangle$ and $\langle u, av \rangle = \bar{a}\langle u, v \rangle$. Many books, especially in physics, put the conjugate in the first slot, so that their $\langle u, v \rangle$ is our $\langle v, u \rangle$. Every formula involving a lone conjugate then flips. When you read a formula elsewhere, check the convention first; when you write one, say which you use.

Quick check. Is $\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + x_2y_2 + x_1y_2$ an inner product on $\mathbb{R}^2$? If not, name the clause that fails and give a witness.

Solution. No. It is bilinear, and $\langle \mathbf{x}, \mathbf{x} \rangle = x_1^2 + x_2^2 + x_1x_2 = (x_1 + \frac{1}{2}x_2)^2 + \frac{3}{4}x_2^2 > 0$ for $\mathbf{x} \neq \mathbf{0}$, so (IP1) and (IP3) both hold. The clause that fails is (IP2): with $\mathbf{u} = \mathbf{e}_1$ and $\mathbf{v} = \mathbf{e}_2$ we get $\langle \mathbf{u}, \mathbf{v} \rangle = 1$ but $\langle \mathbf{v}, \mathbf{u} \rangle = 0$. Symmetrizing the cross term, to $x_1y_1 + x_2y_2 + \frac{1}{2}(x_1y_2 + x_2y_1)$, repairs it.

11.1.5 Norms

By [Theorem 11.1.1](#) (a), the number $\langle v, v \rangle$ is real and ≥ 0 , so it has a unique nonnegative real square root. That square root is the length we were after.

Definition 11.1.3. Induced Norm

Let V be an inner product space. The **norm** of $v \in V$ is

$$\|v\| := \sqrt{\langle v, v \rangle},$$

the **nonnegative** real square root. A vector with $\|v\| = 1$ is a **unit vector**, and replacing a nonzero v by $v/\|v\|$ is called **normalizing** it.

On $\mathbb{R}^n$ with the dot product this is the Euclidean length $\sqrt{x_1^2 + \cdots + x_n^2}$, and on $\mathbb{C}^n$ it is $\sqrt{\sum |x_i|^2}$. On $C[0, 1]$ it is $(\int_0^1 |f|^2)^{1/2}$, and on matrices it is the **Frobenius norm** $(\sum |a_{ij}|^2)^{1/2}$. By definition $\|v\|^2 = \langle v, v \rangle$ always, and this identity, read from right to left, is the most-used step in the chapter: **work with the norm squared, then expand**.

Theorem 11.1.2. Properties of the Induced Norm

Let V be an inner product space over F , $u, v \in V$ and $c \in F$. Then:

- (a) $\|v\| \geq 0$, with $\|v\| = 0$ if and only if $v = 0$;
- (b) $\|cv\| = |c|\|v\|$;
- (c) $\|u + v\|^2 = \|u\|^2 + 2\operatorname{Re}\langle u, v \rangle + \|v\|^2$.

Proof.

- (a) Immediate from [Theorem 11.1.1](#) (a) and the fact that a nonnegative real number has square root 0 only when it is 0.
- (b) By (IP1) and [Theorem 11.1.1](#) (b), $\langle cv, cv \rangle = c\bar{c}\langle v, v \rangle = |c|^2\|v\|^2$ using Theorem 0.6.2 (c). Taking nonnegative square roots gives $\|cv\| = |c|\|v\|$, since $|c| \geq 0$.
- (c) Expand one slot at a time, by (IP1) and then [Theorem 11.1.1](#) (b):

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

By (IP2), $\langle v, u \rangle = \overline{\langle u, v \rangle}$, and $z + \bar{z} = 2 \operatorname{Re} z$ by Theorem 0.6.2 (c). This gives the stated identity. ■

Part (c) is the workhorse. Over $\mathbb{R}$ the middle term is just $2\langle u, v \rangle$, and the identity is the familiar $(a + b)^2 = a^2 + 2ab + b^2$ for vectors. Replacing v by $-v$ gives $\|u - v\|^2 = \|u\|^2 - 2\operatorname{Re}\langle u, v \rangle + \|v\|^2$, and adding the two versions gives the **parallelogram law**

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2. \quad (*)$$

We return to $(*)$ in the last warning of this section.

Warning. **A squared norm is real and nonnegative, never complex.** If a computation in a complex space produces $\|\mathbf{v}\|^2 = 2i$, or a negative number, something has gone wrong — most often a missing conjugate. For $\mathbf{v} = (1, i) \in \mathbb{C}^2$, the correct value is $\|\mathbf{v}\|^2 = |1|^2 + |i|^2 = 2$, while the conjugate-free expression $\mathbf{v}^\top \mathbf{v}$ gives 0. Use the answer as a check on the work.

11.1.6 Orthogonality and Pythagoras

With a norm in hand, the special relation $\langle u, v \rangle = 0$ deserves its name.

Definition 11.1.4. Orthogonal Vectors

Let V be an inner product space. Vectors $u, v \in V$ are **orthogonal**, written $u \perp v$, if $\langle u, v \rangle = 0$. A list $(v_1, \dots, v_k)$ is **orthogonal** if $v_i \perp v_j$ whenever $i \neq j$.

The relation is symmetric: $\langle u, v \rangle = 0$ if and only if $\langle v, u \rangle = \overline{\langle u, v \rangle} = 0$, so “ $u \perp v$ ” needs no order. The zero vector is orthogonal to everything ([Theorem 11.1.1](#) (c)), and by part (a) of the same result it is the **only** vector orthogonal to itself. In $\mathbb{R}^n$ this is perpendicularity in the usual sense, by the law of cosines computation above: $\cos \theta = 0$ exactly when $\mathbf{x} \cdot \mathbf{y} = 0$. In $\mathbb{C}^2$, the vectors $(1, i)$ and $(1, -i)$ are orthogonal, since $1 \cdot \bar{1} + i \cdot \overline{(-i)} = 1 + i \cdot i = 0$.

Theorem 11.1.3. Pythagorean Theorem

Let $(v_1, \dots, v_k)$ be an orthogonal list in an inner product space. Then

$$\|v_1 + \dots + v_k\|^2 = \|v_1\|^2 + \dots + \|v_k\|^2.$$

Proof. Expand the left side by (IP1) and [Theorem 11.1.1](#) (b):

$$\left\langle \sum_i v_i, \sum_j v_j \right\rangle = \sum_i \sum_j \langle v_i, v_j \rangle.$$

Every term with $i \neq j$ is 0, because the list is orthogonal. What survives is $\sum_i \langle v_i, v_i \rangle = \sum_i \|v_i\|^2$, as claimed. ■

So orthogonality is exactly the condition under which squared lengths add. This is why the relation is worth naming, and why the rest of the chapter works so hard to produce orthogonal vectors.

Warning. Over $\mathbb{C}$, the converse of Pythagoras fails. By [Theorem 11.1.2](#) (c), $\|u+v\|^2 = \|u\|^2 + \|v\|^2$ says only that $\operatorname{Re}\langle u, v \rangle = 0$, not that $\langle u, v \rangle = 0$. In $\mathbb{C}$ with $\langle z, w \rangle = z\bar{w}$, take $u = 1$ and $v = i$: then $\langle u, v \rangle = \bar{i} = -i \neq 0$, yet $\|1+i\|^2 = 2 = \|1\|^2 + \|i\|^2$. Over $\mathbb{R}$ the converse does hold, since there Re does nothing.

Theorem 11.1.4. Orthogonal Nonzero Vectors Are Independent

Let $(v_1, \dots, v_k)$ be an orthogonal list of **nonzero** vectors in an inner product space. Then the list is linearly independent.

Idea. Start a dependence relation in the standard way, and then do the one thing an inner product allows: pair the relation with a single v_j . Orthogonality kills every term except the j -th, and what is left is $a_j\|v_j\|^2 = 0$ with $\|v_j\| \neq 0$.

Proof. Let $a_1v_1 + \dots + a_kv_k = 0$ with $a_1, \dots, a_k \in F$. Fix j and take the inner product of both sides with v_j . By (IP1) and [Theorem 11.1.1](#) (c),

$$0 = \left\langle \sum_i a_i v_i, v_j \right\rangle = \sum_i a_i \langle v_i, v_j \rangle = a_j \|v_j\|^2,$$

where the last equality holds because $\langle v_i, v_j \rangle = 0$ for $i \neq j$. Since $v_j \neq 0$, [Theorem 11.1.1](#) (a) gives $\|v_j\|^2 > 0$, so $a_j = 0$. As j was arbitrary, all the coefficients vanish, and the list is linearly independent by Definition 1.4.1. ■

The hypothesis “nonzero” cannot be dropped: the list $(\mathbf{e}_1, \mathbf{0})$ is orthogonal and dependent. With it, the result is a free supply of independent lists, and this is what makes the next section

possible.

11.1.7 The Cauchy–Schwarz inequality

Angles came from the formula $\cos \theta = (\mathbf{x} \cdot \mathbf{y})/(\|\mathbf{x}\|\|\mathbf{y}\|)$. To define an angle by that formula in an abstract space, we must know that the right-hand side lies in $[-1, 1]$. That is exactly the following inequality, the most used inequality in the subject.

Theorem 11.1.5. Cauchy–Schwarz Inequality

Let V be an inner product space and $u, v \in V$. Then

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

with equality if and only if the list (u, v) is linearly dependent.

Idea. Picture u and v in $\mathbb{R}^2$. Drop a perpendicular from the tip of u onto the line through v . The foot of the perpendicular is a multiple cv , and the two legs cv and $u - cv$ of the resulting right triangle have squared lengths adding to $\|u\|^2$. Since one leg has squared length $|\langle u, v \rangle|^2/\|v\|^2$, the inequality is just “a leg is no longer than the hypotenuse”, and equality means the other leg is 0. The picture proves nothing in an abstract space, so we check algebraically that $u - cv \perp v$ for the right c , and then let Pythagoras do the work.

Proof. If $v = 0$, both sides are 0 by [Theorem 11.1.1](#) (c), and $(u, 0)$ is dependent, so the statement holds. Suppose $v \neq 0$, so $\|v\| > 0$, and put

$$c = \frac{\langle u, v \rangle}{\|v\|^2}, \quad z = u - cv.$$

Then $z \perp v$, since by (IP1)

$$\langle z, v \rangle = \langle u, v \rangle - c\langle v, v \rangle = \langle u, v \rangle - \langle u, v \rangle = 0,$$

and hence also $z \perp cv$. Now $u = z + cv$ is a sum of two orthogonal vectors, so [Theorem 11.1.3](#) and [Theorem 11.1.2](#) (b) give

$$\|u\|^2 = \|z\|^2 + |c|^2\|v\|^2 = \|z\|^2 + \frac{|\langle u, v \rangle|^2}{\|v\|^2}. \quad (\dagger)$$

Since $\|z\|^2 \geq 0$, dropping it gives $\|u\|^2\|v\|^2 \geq |\langle u, v \rangle|^2$, and taking nonnegative square roots proves the inequality.

Equality holds in $(\dagger)$ exactly when $\|z\| = 0$, that is ([Theorem 11.1.2](#) (a)) exactly when $u = cv$. If $u = cv$, the list (u, v) is dependent. Conversely, suppose (u, v) is dependent, say $au + bv = 0$ with a, b not both 0. Here $a \neq 0$, since $a = 0$ would give $bv = 0$ with $b \neq 0$ and hence $v = 0$, which we have excluded. So $u = dv$ with $d = -b/a$, and then $c = d\langle v, v \rangle/\|v\|^2 = d$, so $z = 0$. This proves the equality case. ■

The proof is worth rereading for its shape: the vector z was **found** from a picture and then **checked** from the axioms. That is the standard way to work in an abstract inner product space.

Corollary 11.1.6. Triangle Inequality

Let V be an inner product space and $u, v \in V$. Then

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

Proof. For any complex number z we have $\operatorname{Re} z \leq |z|$, as in the proof of Theorem 0.6.4. Hence, by Theorem 11.1.2 (c) and then Theorem 11.1.5,

$$\begin{aligned} \|u + v\|^2 &= \|u\|^2 + 2 \operatorname{Re} \langle u, v \rangle + \|v\|^2 \\ &\leq \|u\|^2 + 2|\langle u, v \rangle| + \|v\|^2 \\ &\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\ &= (\|u\| + \|v\|)^2. \end{aligned}$$

Both $\|u + v\|$ and $\|u\| + \|v\|$ are ≥ 0 , so taking square roots preserves the inequality. ■

Together, Theorem 11.1.2 (a), (b), and Corollary 11.1.6 say that $\|\cdot\|$ is a **norm** in the general sense studied in Chapter 16. Chapter 16 also shows that not every norm arises from an inner product; the last warning of this section explains how to tell.

Definition 11.1.5. Angle Between Vectors

Let V be a **real** inner product space and let $u, v \in V$ be **nonzero**. The **angle** between u and v is the unique $\theta \in [0, \pi]$ with

$$\cos \theta = \frac{\langle u, v \rangle}{\|u\| \|v\|}.$$

The definition needs Theorem 11.1.5 twice over. Cauchy–Schwarz says the quotient has absolute value at most 1, so it lies in $[-1, 1]$; and $\cos$ is a strictly decreasing bijection from $[0, \pi]$ onto $[-1, 1]$, so exactly one θ qualifies. Without the inequality the formula could ask for the angle whose cosine is 7. Note also that $\theta = \pi/2$ if and only if $\langle u, v \rangle = 0$, so the definition is consistent with Definition 11.1.4.

Over $\mathbb{C}$ there is no angle, because $\langle u, v \rangle$ is not real. The quantity $|\langle u, v \rangle|/(\|u\|\|v\|) \in [0, 1]$ still makes sense, and it still measures how far the two vectors are from being orthogonal.

Example. Three Angles

Compute the angle between $\mathbf{u} = (1, 1, 0)$ and $\mathbf{v} = (1, 0, 1)$ in $\mathbb{R}^3$ with the dot product; between $p = 1$ and $q = x$ in $\mathbb{R}[x]_{\leq 1}$ with $\langle p, q \rangle = \int_0^1 pq$; and between $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and

$\mathbf{B} = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$ in $M_2(\mathbb{R})$ with the Frobenius inner product.

Solution. In $\mathbb{R}^3$: $\langle \mathbf{u}, \mathbf{v} \rangle = 1$, $\|\mathbf{u}\| = \|\mathbf{v}\| = \sqrt{2}$, so $\cos \theta = 1/2$ and $\theta = \pi/3$.

In $\mathbb{R}[x]_{\leq 1}$: $\langle 1, x \rangle = \int_0^1 t \, dt = \frac{1}{2}$, $\|1\|^2 = \int_0^1 1 \, dt = 1$, and $\|x\|^2 = \int_0^1 t^2 \, dt = \frac{1}{3}$. Hence

$$\cos \theta = \frac{1/2}{1 \cdot 1/\sqrt{3}} = \frac{\sqrt{3}}{2},$$

so $\theta = \pi/6$. The constant 1 and the function x are far from orthogonal on $[0, 1]$, which is believable: both are positive there.

In $M_2(\mathbb{R})$: $\langle \mathbf{A}, \mathbf{B} \rangle = 1 - 1 + 1 - 1 = 0$, so $\theta = \pi/2$. The two matrices are orthogonal, with $\|\mathbf{A}\| = \|\mathbf{B}\| = 2$. As a check on Pythagoras, $\mathbf{A} + \mathbf{B} = \begin{pmatrix} 2 & 0 \\ 2 & 0 \end{pmatrix}$ has $\|\mathbf{A} + \mathbf{B}\|^2 = 8 = 4 + 4$.

Quick check. In $\mathbb{C}^2$ with the standard inner product, compute $\langle (1, i), (1, 1) \rangle$ and $\langle (1, 1), (1, i) \rangle$. Check the answers against (IP2), and against Cauchy–Schwarz.

Solution. $\langle (1, i), (1, 1) \rangle = 1 \cdot \bar{1} + i \cdot \bar{1} = 1 + i$, and $\langle (1, 1), (1, i) \rangle = 1 \cdot \bar{1} + 1 \cdot \bar{i} = 1 - i$. They are conjugates, as (IP2) requires. For Cauchy–Schwarz, $|1 + i| = \sqrt{2}$, while $\|(1, i)\| = \sqrt{2}$ and $\|(1, 1)\| = \sqrt{2}$, so the inequality reads $\sqrt{2} \leq 2$. It is strict, which is right: $(1, i)$ and $(1, 1)$ are independent.

Warning. Not every norm comes from an inner product. A function $\|\cdot\|$ can satisfy positivity, homogeneity, and the inequality of Corollary 11.1.6 without being $\sqrt{\langle v, v \rangle}$ for any inner product. The test is the parallelogram law (*): every induced norm satisfies it, as we derived from Theorem 11.1.2 (c), so a norm that violates it is induced by nothing. Take $\|\mathbf{x}\|_\infty = \max(|x_1|, |x_2|)$ on $\mathbb{R}^2$ with $\mathbf{u} = \mathbf{e}_1$, $\mathbf{v} = \mathbf{e}_2$. Then $\|\mathbf{u} + \mathbf{v}\|_\infty = \|\mathbf{u} - \mathbf{v}\|_\infty = 1$, so the left side of (*) is 2, while the right side is 4. Remarkably, the law is not only necessary but sufficient; Exercise (C1: The parallelogram law characterizes inner-product norms) proves the substance of that in $\mathbb{R}^2$, and Chapter 16 takes up general norms.

11.1.8 Exercises

11.1.8.1 A. Check your understanding

Exercise. A1

- State the three axioms of an inner product on a vector space over $F = \mathbb{R}$ or $\mathbb{C}$, and say in which slot the map is linear.
- Why must the field be $\mathbb{R}$ or $\mathbb{C}$? Name the two features used.

- (c) Explain why $\langle v, v \rangle$ is always a real number.
- (d) True or false: an inner product on a complex space is linear in the second slot. Justify your answer.
- (e) State the Cauchy–Schwarz inequality, its equality case, and one place where the equality case is used.
- (f) True or false: if $\|u + v\|^2 = \|u\|^2 + \|v\|^2$ in a complex inner product space, then $u \perp v$. Justify your answer.

Solution.

- (a) (IP1) $\langle au + bv, w \rangle = a\langle u, w \rangle + b\langle v, w \rangle$; (IP2) $\langle v, u \rangle = \overline{\langle u, v \rangle}$; (IP3) $\langle v, v \rangle > 0$ for $v \neq 0$ (Definition 11.1.1). It is linear in the **first** slot.
- (b) Positivity, so that $\langle v, v \rangle > 0$ has meaning and square roots exist, and conjugation, so that $\langle v, v \rangle$ is real and nonnegative over $\mathbb{C}$.
- (c) By (IP2) with $u = v$, the number equals its own conjugate, so its imaginary part is 0 (Theorem 11.1.1 (a)).
- (d) False. It is **conjugate**-linear in the second slot: $\langle u, av \rangle = \bar{a}\langle u, v \rangle$ (Theorem 11.1.1 (b)). For example, in $\mathbb{C}$ with $\langle z, w \rangle = z\bar{w}$, $\langle 1, i \rangle = -i$ while $i\langle 1, 1 \rangle = i$.
- (e) $|\langle u, v \rangle| \leq \|u\|\|v\|$, with equality if and only if (u, v) is dependent (Theorem 11.1.5). The inequality is what makes Definition 11.1.5 well posed: it puts the quotient $\langle u, v \rangle / (\|u\|\|v\|)$ inside $[-1, 1]$, where arccos is defined. The equality case then says which vectors sit at the ends of that interval: $\cos \theta = \pm 1$ exactly for dependent pairs, that is, $\theta = 0$ or $\theta = \pi$.
- (f) False. It says only $\operatorname{Re}\langle u, v \rangle = 0$. In $\mathbb{C}$ with $\langle z, w \rangle = z\bar{w}$, take $u = 1$, $v = i$: then $\langle u, v \rangle = -i \neq 0$ but $\|1 + i\|^2 = 2 = 1 + 1$.

11.1.8.2 B. Practice

Exercise. B1: Determine which are inner products

Determine which of the following are inner products on the stated space. For those that are, verify the axioms; for those that are not, name the clause that fails and give an explicit witness.

- (a) On $\mathbb{R}^2$: $\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + 3x_2y_2$.
- (b) On $\mathbb{R}^2$: $\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2$.
- (c) On $\mathbb{C}^2$: $\langle \mathbf{x}, \mathbf{y} \rangle = x_1\bar{y}_1 + x_2y_2$.
- (d) On $M_2(\mathbb{R})$: $\langle \mathbf{A}, \mathbf{B} \rangle = \operatorname{tr}(\mathbf{AB})$.
- (e) On $\mathbb{R}[x]_{\leq 2}$: $\langle p, q \rangle = p(0)q(0) + p(1)q(1) + p(2)q(2)$.

Solution.

- (a) An inner product: it is the weighted product with weights $1, 3 > 0$. Bilinearity and symmetry are clear, and $\langle \mathbf{x}, \mathbf{x} \rangle = x_1^2 + 3x_2^2 = 0$ forces $x_1 = x_2 = 0$.
- (b) An inner product. It is bilinear term by term, and symmetric because the two cross

terms are interchanged by swapping $\mathbf{x}$ and $\mathbf{y}$. For positivity, complete the square:

$$\langle \mathbf{x}, \mathbf{x} \rangle = x_1^2 + 2x_1x_2 + 2x_2^2 = (x_1 + x_2)^2 + x_2^2,$$

which is 0 only if $x_2 = 0$ and $x_1 + x_2 = 0$, that is $\mathbf{x} = \mathbf{0}$.

- (c) Not an inner product. (IP2) fails: with $\mathbf{x} = (0, 1)$ and $\mathbf{y} = (0, i)$ we get $\langle \mathbf{x}, \mathbf{y} \rangle = i$ and $\langle \mathbf{y}, \mathbf{x} \rangle = i$, while (IP2) would require $\langle \mathbf{y}, \mathbf{x} \rangle = \bar{i} = -i$. Positivity fails too, at $\mathbf{x} = (0, i)$: $\langle \mathbf{x}, \mathbf{x} \rangle = i^2 = -1$.
- (d) Not an inner product. It is bilinear, and symmetric by Theorem 0.9.6 (3). But (IP3) fails for $\mathbf{N} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \mathbf{O}$, since $\mathbf{N}^2 = \mathbf{O}$ and so $\langle \mathbf{N}, \mathbf{N} \rangle = \text{tr}(\mathbf{N}^2) = 0$. Replacing $\text{tr}(\mathbf{AB})$ by $\text{tr}(\mathbf{B}^\top \mathbf{A})$ repairs it: that is the Frobenius inner product.
- (e) An inner product: the node example with the three distinct nodes 0, 1, 2 on $\mathbb{R}[x]_{\leq 2}$. Bilinearity and symmetry come from linearity of evaluation. If $\langle p, p \rangle = p(0)^2 + p(1)^2 + p(2)^2 = 0$, then p has three distinct roots and degree at most 2, so $p = 0$ by Corollary 6.1.4.

Exercise. B2: Norms and angles

- (a) In $\mathbb{R}^3$ with the dot product, find $\|\mathbf{u}\|$, $\|\mathbf{v}\|$, and the cosine of the angle between $\mathbf{u} = (1, 2, 2)$ and $\mathbf{v} = (2, -1, 2)$.
- (b) In $\mathbb{R}[x]_{\leq 2}$ with $\langle p, q \rangle = \int_{-1}^1 pq$, compute $\|1\|$, $\|x\|$, and $\langle 1, x \rangle$. What is the angle between 1 and x ?
- (c) In $M_2(\mathbb{C})$ with the Frobenius inner product, compute $\|\mathbf{A}\|$ for $\mathbf{A} = \begin{pmatrix} 1 & i \\ 0 & 1+i \end{pmatrix}$.

Solution.

- (a) $\|\mathbf{u}\| = \sqrt{1+4+4} = 3$ and $\|\mathbf{v}\| = \sqrt{4+1+4} = 3$. Also $\langle \mathbf{u}, \mathbf{v} \rangle = 2 - 2 + 4 = 4$, so $\cos \theta = 4/9$. Cauchy-Schwarz reads $4 \leq 9$, and it is strict, as it must be for an independent pair.
- (b) $\|1\|^2 = \int_{-1}^1 1 \, dt = 2$, so $\|1\| = \sqrt{2}$; $\|x\|^2 = \int_{-1}^1 t^2 \, dt = \frac{2}{3}$, so $\|x\| = \sqrt{2/3}$; and $\langle 1, x \rangle = \int_{-1}^1 t \, dt = 0$. The angle is $\pi/2$: on a symmetric interval, the constants are orthogonal to the odd functions. Compare this with [Example \(Three Angles\)](#), where the interval $[0, 1]$ gave the angle $\pi/6$ instead. **The geometry depends on the inner product, not only on the vectors.**
- (c) $\|\mathbf{A}\|^2 = |1|^2 + |i|^2 + 0 + |1+i|^2 = 1 + 1 + 2 = 4$, so $\|\mathbf{A}\| = 2$.

Exercise. B3: Cauchy-Schwarz, concretely

- (a) Verify [Theorem 11.1.5](#) for $\mathbf{a} = (1, 2, 3)$ and $\mathbf{b} = (3, 2, 1)$ in $\mathbb{R}^3$ with the dot product, and say whether the inequality is strict.

(b) Deduce from [Theorem 11.1.5](#) the classical inequality

$$\left(\sum_{i=1}^n a_i b_i\right)^2 \leq \left(\sum_{i=1}^n a_i^2\right) \left(\sum_{i=1}^n b_i^2\right)$$

for all real a_i, b_i .

(c) Hence show that $(a_1 + \cdots + a_n)^2 \leq n(a_1^2 + \cdots + a_n^2)$ for all real a_i .

Solution.

- (a) $\langle \mathbf{a}, \mathbf{b} \rangle = 3 + 4 + 3 = 10$, so $|\langle \mathbf{a}, \mathbf{b} \rangle|^2 = 100$, while $\|\mathbf{a}\|^2 \|\mathbf{b}\|^2 = 14 \cdot 14 = 196$. So $10 \leq 14$, strictly. It must be strict, since $\mathbf{a}$ and $\mathbf{b}$ are not multiples of each other.
- (b) Apply [Theorem 11.1.5](#) on $\mathbb{R}^n$ with the dot product to $\mathbf{a} = (a_1, \dots, a_n)$ and $\mathbf{b} = (b_1, \dots, b_n)$. Then $\langle \mathbf{a}, \mathbf{b} \rangle = \sum a_i b_i$, $\|\mathbf{a}\|^2 = \sum a_i^2$, and $\|\mathbf{b}\|^2 = \sum b_i^2$, and squaring the inequality $|\langle \mathbf{a}, \mathbf{b} \rangle| \leq \|\mathbf{a}\| \|\mathbf{b}\|$ gives the claim.
- (c) Take $b_i = 1$ for all i in (b). Then $\sum a_i b_i = \sum a_i$ and $\sum b_i^2 = n$, so $(\sum a_i)^2 \leq n \sum a_i^2$. Equality holds exactly when $\mathbf{a}$ is a multiple of $(1, \dots, 1)$, that is, when all the a_i are equal.

Exercise. B4: A lower bound for a difference of norms

Let V be an inner product space and $u, v \in V$. Prove that

$$|\|u\| - \|v\|| \leq \|u - v\|.$$

Solution. Write $u = (u - v) + v$. By [Corollary 11.1.6](#),

$$\|u\| \leq \|u - v\| + \|v\|,$$

so $\|u\| - \|v\| \leq \|u - v\|$. Swapping the roles of u and v gives $\|v\| - \|u\| \leq \|v - u\|$, and $\|v - u\| = |-1|\|u - v\| = \|u - v\|$ by [Theorem 11.1.2](#) (b). The quantity $|\|u\| - \|v\||$ is one of the two numbers just bounded, so it is at most $\|u - v\|$. This proves the inequality.

11.1.8.3 C. Going deeper

Exercise. C1: The parallelogram law characterizes inner-product norms

Let $\|\cdot\|$ be a norm on $\mathbb{R}^2$, that is, a function $\mathbb{R}^2 \rightarrow \mathbb{R}$ satisfying the three properties of [Theorem 11.1.2](#) (a), (b), and [Corollary 11.1.6](#). Suppose it satisfies the parallelogram law

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$$

for all $\mathbf{u}, \mathbf{v}$, and define $\beta(\mathbf{u}, \mathbf{v}) = \frac{1}{4}(\|\mathbf{u} + \mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2)$.

- (a) Show that $\beta(\mathbf{u}, \mathbf{u}) = \|\mathbf{u}\|^2$, that $\beta(\mathbf{u}, \mathbf{v}) = \beta(\mathbf{v}, \mathbf{u})$, and that $\beta(\mathbf{u}, \mathbf{0}) = 0$.
- (b) Prove that $\beta(\mathbf{u}, \mathbf{v}) + \beta(\mathbf{u}, \mathbf{w}) = 2\beta(\mathbf{u}, \frac{1}{2}(\mathbf{v} + \mathbf{w}))$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w}$.
- (c) Deduce that $\beta(\mathbf{u}, \mathbf{v}) + \beta(\mathbf{u}, \mathbf{w}) = \beta(\mathbf{u}, \mathbf{v} + \mathbf{w})$.

- (d) Granting that $\beta(\mathbf{u}, t\mathbf{v}) = t\beta(\mathbf{u}, \mathbf{v})$ for all $t \in \mathbb{R}$, conclude that β is an inner product on $\mathbb{R}^2$ whose induced norm is $\|\cdot\|$.

Hint: for (b), apply the parallelogram law to the pair $\mathbf{u} + \frac{1}{2}(\mathbf{v} + \mathbf{w})$, $\frac{1}{2}(\mathbf{v} - \mathbf{w})$, and then to the pair $\mathbf{u} - \frac{1}{2}(\mathbf{v} + \mathbf{w})$, $\frac{1}{2}(\mathbf{v} - \mathbf{w})$.

Solution.

- (a) $\beta(\mathbf{u}, \mathbf{u}) = \frac{1}{4}(\|2\mathbf{u}\|^2 - \|\mathbf{0}\|^2) = \frac{1}{4} \cdot 4\|\mathbf{u}\|^2 = \|\mathbf{u}\|^2$, using homogeneity. Symmetry holds because $\|\mathbf{v} - \mathbf{u}\| = \|\mathbf{u} - \mathbf{v}\|$, again by homogeneity with $c = -1$. Finally $\beta(\mathbf{u}, \mathbf{0}) = \frac{1}{4}(\|\mathbf{u}\|^2 - \|\mathbf{u}\|^2) = 0$.
- (b) Put $\mathbf{m} = \frac{1}{2}(\mathbf{v} + \mathbf{w})$ and $\mathbf{d} = \frac{1}{2}(\mathbf{v} - \mathbf{w})$, so that $\mathbf{v} = \mathbf{m} + \mathbf{d}$ and $\mathbf{w} = \mathbf{m} - \mathbf{d}$. Apply the parallelogram law to $\mathbf{u} + \mathbf{m}$ and $\mathbf{d}$: since $(\mathbf{u} + \mathbf{m}) + \mathbf{d} = \mathbf{u} + \mathbf{v}$ and $(\mathbf{u} + \mathbf{m}) - \mathbf{d} = \mathbf{u} + \mathbf{w}$,

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

Apply it to $\mathbf{u} - \mathbf{m}$ and $\mathbf{d}$: since $(\mathbf{u} - \mathbf{m}) - \mathbf{d} = \mathbf{u} - \mathbf{v}$ and $(\mathbf{u} - \mathbf{m}) + \mathbf{d} = \mathbf{u} - \mathbf{w}$,

$$\|\mathbf{u} - \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{w}\|^2 = 2\|\mathbf{u} - \mathbf{m}\|^2 + 2\|\mathbf{d}\|^2.$$

Subtracting the second from the first, the terms $2\|\mathbf{d}\|^2$ cancel, and dividing by 4 gives $\beta(\mathbf{u}, \mathbf{v}) + \beta(\mathbf{u}, \mathbf{w}) = 2\beta(\mathbf{u}, \mathbf{m})$, which is the claim.

- (c) Take $\mathbf{w} = \mathbf{0}$ in (b). By (a), $\beta(\mathbf{u}, \mathbf{0}) = 0$, so $\beta(\mathbf{u}, \mathbf{v}) = 2\beta(\mathbf{u}, \frac{1}{2}\mathbf{v})$ for all $\mathbf{v}$. Applying this with $\mathbf{v} + \mathbf{w}$ in place of $\mathbf{v}$, we get $2\beta(\mathbf{u}, \frac{1}{2}(\mathbf{v} + \mathbf{w})) = \beta(\mathbf{u}, \mathbf{v} + \mathbf{w})$, and the left side is $\beta(\mathbf{u}, \mathbf{v}) + \beta(\mathbf{u}, \mathbf{w})$ by (b).
- (d) By (c) and the granted homogeneity, β is linear in its second slot, and by the symmetry in (a) it is then linear in the first as well, so (IP1) and (IP2) hold (conjugation is trivial over $\mathbb{R}$). For (IP3), $\beta(\mathbf{u}, \mathbf{u}) = \|\mathbf{u}\|^2 > 0$ for $\mathbf{u} \neq \mathbf{0}$, since a norm vanishes only at $\mathbf{0}$. So β is an inner product, and its induced norm is $\sqrt{\beta(\mathbf{u}, \mathbf{u})} = \|\mathbf{u}\|$.

Remark. The granted step in (d) is where analysis enters: additivity from (c) gives $\beta(\mathbf{u}, t\mathbf{v}) = t\beta(\mathbf{u}, \mathbf{v})$ for every **rational** t by induction, and the extension to real t uses that $t \mapsto \beta(\mathbf{u}, t\mathbf{v})$ is continuous, which follows from continuity of the norm. Nothing else in the argument needs anything beyond the three norm properties.

Exercise. C2: Recovering the inner product from the norm

Let V be an inner product space over F .

- (a) For $F = \mathbb{R}$, prove the polarization identity $\langle u, v \rangle = \frac{1}{4}(\|u + v\|^2 - \|u - v\|^2)$.
- (b) For $F = \mathbb{C}$, prove that

$$\langle u, v \rangle = \frac{1}{4} \sum_{k=0}^3 i^k \|u + i^k v\|^2.$$

- (c) Deduce that two inner products on V with the same induced norm are equal.

Solution.

- (a) By [Theorem 11.1.2](#) (c) and its version with $-v$, and since Re is the identity on $\mathbb{R}$,

$$\|u + v\|^2 - \|u - v\|^2 = 4\langle u, v \rangle.$$

Dividing by 4 gives the identity.

- (b) Write $z = \langle u, v \rangle$. By [Theorem 11.1.2](#) (c) and [Theorem 11.1.1](#) (b),

$$\|u + i^k v\|^2 = \|u\|^2 + \|v\|^2 + 2\text{Re}(\overline{i^k} z),$$

where we used $\|i^k v\| = |i^k| \|v\| = \|v\|$. Now $\sum_{k=0}^3 i^k = 1 + i - 1 - i = 0$, so the terms $\|u\|^2 + \|v\|^2$ contribute nothing. Writing $z = a + bi$ with $a, b \in \mathbb{R}$, the four remaining terms $2i^k \text{Re}(\overline{i^k} z)$ are, for $k = 0, 1, 2, 3$,

$$2a, \quad 2ib, \quad 2a, \quad 2ib,$$

since $\text{Re}(z) = a$, $\text{Re}(-iz) = b$, $\text{Re}(-z) = -a$ and $\text{Re}(iz) = -b$. Their sum is $4a + 4bi = 4z$, and dividing by 4 gives the identity.

- (c) Both (a) and (b) express $\langle u, v \rangle$ using only the norm. So if two inner products induce the same norm, the right-hand sides agree for all u, v , hence so do the left-hand sides.

Exercise. C3: The equality cases

Let V be an inner product space.

- Prove that $\|u + v\| = \|u\| + \|v\|$ if and only if one of u, v is a **nonnegative real** multiple of the other.
- Deduce the equality condition in the classical inequality of [Exercise \(B3: Cauchy–Schwarz, concretely\)](#) (b): for real a_i, b_i , equality holds if and only if one of the vectors $(a_1, \dots, a_n)$, $(b_1, \dots, b_n)$ is a multiple of the other.
- Give real numbers a_i, b_i , not all zero, with $\sum a_i b_i < 0$ and equality in the classical inequality, and explain why this does **not** contradict (a).

Solution.

- ($\Leftarrow$) If $v = 0$ both sides are $\|u\|$. If $u = tv$ with $t \geq 0$, then $\|u + v\| = (t + 1)\|v\| = t\|v\| + \|v\| = \|u\| + \|v\|$ by [Theorem 11.1.2](#) (b).
($\Rightarrow$) Suppose $\|u + v\| = \|u\| + \|v\|$. Squaring and using [Theorem 11.1.2](#) (c),

$$2\text{Re}\langle u, v \rangle = 2\|u\|\|v\|.$$

The chain in the proof of [Corollary 11.1.6](#) is therefore an equality throughout, so $\text{Re}\langle u, v \rangle = |\langle u, v \rangle| = \|u\|\|v\|$. The second equality and [Theorem 11.1.5](#) make (u, v) dependent. Assume $v \neq 0$ (otherwise $v = 0u$ and we are done) and write $u = cv$. Then $\langle u, v \rangle = c\|v\|^2$, and the first equality says $\text{Re}(c)\|v\|^2 = |c|\|v\|^2$, so $\text{Re} c = |c|$, which forces c to be real and ≥ 0 .

- Apply [Theorem 11.1.5](#) on $\mathbb{R}^n$ as in [Exercise \(B3: Cauchy–Schwarz, concretely\)](#) (b). Squaring removes the sign, so equality in the squared inequality is equality in $|\langle \mathbf{a}, \mathbf{b} \rangle| = \|\mathbf{a}\|\|\mathbf{b}\|$, which by [Theorem 11.1.5](#) holds exactly when $(\mathbf{a}, \mathbf{b})$ is dependent, that is, when one of the two vectors is a multiple of the other.
- Take $n = 1$, $a_1 = 1$, $b_1 = -1$: then $\sum a_i b_i = -1$ and both sides of the classical

inequality equal 1. There is no contradiction with (a), because (a) is about equality in [Corollary 11.1.6](#), which needs a nonnegative multiple, while the classical inequality is the **squared** Cauchy–Schwarz inequality, which is blind to the sign. Here $\|\mathbf{a} + \mathbf{b}\| = 0 \neq 2 = \|\mathbf{a}\| + \|\mathbf{b}\|$.