

Lecture 2

Rigid geometry using inner products

2.1 Inner product spaces

Inner products generalize rigid Euclidean geometry to \mathbb{C}^n .

Definition 8. An inner product on a vector space V over \mathbb{C} is a function that assigns to every pair of vectors x, y a scalar $\langle x, y \rangle$ satisfying the following for all $x, y, z \in V$ and $\alpha \in \mathbb{C}$:

1. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$. (additive in the first argument)
2. $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$. (scaling in the second argument)
3. $\overline{\langle x, y \rangle} = \langle y, x \rangle$. (conjugate symmetry)
4. $\langle x, x \rangle$ is real and nonnegative if x is not the zero vector. (positive-definite)

Definition 9. An inner product space is a vector space over \mathbb{C} equipped with an inner product.

Theorem 5. Let V be an inner product space. Then for $x, y, z \in V$ and $\alpha \in \mathbb{C}$, the following are true:

1. $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$.
2. $\langle x, cy \rangle = \bar{c} \langle x, y \rangle$.
3. $\langle x, 0 \rangle = \langle 0, x \rangle = 0$.
4. $\langle x, x \rangle = 0$ if and only if $x = 0$.
5. If $\langle x, y \rangle = \langle x, z \rangle$ for all x , then $y = z$.

Very important example: standard inner product on \mathbb{C}^n

Define an inner product on \mathbb{C}^n by $\langle x, y \rangle = y^* x = \sum_{i=1}^n x_i \bar{y}_i$. Let's verify that this proposed inner product satisfies the inner product axioms. It is linear (preserves addition and scaling) in the first argument:

$$\begin{aligned} \langle \alpha x + y, z \rangle &= \sum_{i=1}^n (\alpha x + y)_i \bar{z}_i = \sum_{i=1}^n (\alpha x_i + y_i) \bar{z}_i \\ &= \sum_{i=1}^n (\alpha x_i \bar{z}_i + y_i \bar{z}_i) = \alpha \sum_{i=1}^n x_i \bar{z}_i + \sum_{i=1}^n y_i \bar{z}_i = \alpha \langle x, z \rangle + \langle y, z \rangle. \end{aligned}$$

It is conjugate symmetric:

$$\begin{aligned}\overline{\langle x, y \rangle} &= \overline{\sum_{i=1}^n x_i \overline{y_i}} = \sum_{i=1}^n \overline{x_i \overline{y_i}} \\ &= \sum_{i=1}^n \overline{x_i} \overline{\overline{y_i}} = \sum_{i=1}^n \overline{x_i} y_i = \langle y, x \rangle.\end{aligned}$$

And it’s positive-definite, as

$$\langle x, x \rangle = \sum_{i=1}^n x_i \overline{x_i}, \quad (2.1)$$

which is a sum in which each term is nonnegative real. Therefore if the whole sum equals zero, every term is zero, and x is the zero vector.

This inner product we just defined is called the *standard inner product*, and we will henceforth marry it to \mathbb{C}^n unless otherwise specified. Occasionally, in notationally reckless contexts, we will write it as $y^* x$.

Definition 10 (Length of a vector). *The **length** of a vector v , denoted by $\|v\|$, is $\sqrt{\langle v, v \rangle}$.*

The length of a vector scales with the vector. That is, $\|\alpha v\| = |\alpha| \|v\|$.

2.2 A “law of cosines”

In Euclidean geometry, the Law of Cosines states for a triangle ABC that $a^2 + b^2 - 2ab \cos \theta = c^2$, where θ is the angle at corner C . We will rediscover this fact in \mathbb{C}^n using inner products.

Let x be the displacement vector from one corner of a triangle to the second, and y the displacement vector from the first to the third. Then the three sides of this triangle are x , y , and $x - y$. An identity for $\|x - y\|^2$ follows.

$$\begin{aligned}\|x - y\|^2 &= \langle x - y, x - y \rangle \\ &= \langle x, x \rangle - \langle x, y \rangle - \langle y, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + \|y\|^2 - \langle x, y \rangle - \overline{\langle x, y \rangle} \\ &= \|x\|^2 + \|y\|^2 - 2 \operatorname{Re}\{\langle x, y \rangle\} \\ &= \|x\|^2 + \|y\|^2 - 2 \|x\| \|y\| \left[\frac{\operatorname{Re}\{\langle x, y \rangle\}}{\|x\| \|y\|} \right]. \quad (\text{if neither } x \text{ nor } y \text{ is zero})\end{aligned}$$

The quotient in square brackets should be understood as $\cos \theta$, where $0 \leq \theta < \pi$ is the angle between the vectors x and y .¹ If it is zero, then our identity reduces to the Pythagorean Theorem and we have a right triangle.

Definition 11 (Orthogonal and orthonormal subsets). *A set of vectors $\{v_i\}_{i \in I}$ is **orthogonal** if $\langle v_i, v_j \rangle = 0$ for $i \neq j$. It is **orthonormal** if in addition $\langle v_i, v_i \rangle = 1$.*

Theorem 6 (Pythagoras). *If $\{v_1, v_2, \dots, v_n\}$ is orthogonal, then $\sum_{i=1}^n \|v_i\|^2 = \|\sum_{i=1}^n v_i\|^2$.*

Proof. Split the sum into head and tail, then use induction. □

¹In statistics, this is called Pearson’s correlation coefficient.

Πυθαγόρας didn't know what he was getting himself into! This simple result is richly applicable well beyond geometry class. An electrical interpretation is that when two sinusoidal power sources of different frequencies are superimposed, the resulting power is the sum. Likewise, when two musical instruments play at the same time, the power of the pressure wave is the sum of the instruments' respective powers. (This does not, however, mean that the result is twice as loud, as loudness perception is roughly logarithmic.) You may recall from a statistics class that variances of independent random variables add when the random variables are added. This is because independent random variables have a correlation (cosine) of zero, and are therefore susceptible to the Pythagorean theorem.

2.3 An inequality that bounds sums

Definition 12 (Orthogonal projection). Let $u, v \in V$, with $v \neq 0$. The *orthogonal projection* of u onto v is $\text{proj}_v u = \frac{\langle u, v \rangle}{\langle v, v \rangle} v$.

The motivating property of $\text{proj}_v u$ is that $u - \text{proj}_v u$ is orthogonal to v . proj_v is linear and does not change when v is multiplied by a nonzero scalar.

The following fact asserts that cosines lie between -1 and 1 .

Theorem 7 (Cauchy-(Bunjakowsky)-Schwarz). For all pairs $u, v \in V$, $|\langle u, v \rangle| \leq \|u\| \|v\|$.

Proof. The proof of the Cauchy-Schwarz inequality² is notorious for employing unpredictable magic tricks whence the result comes out of nowhere. I've tried to structure it in a more geometrically luminous, even if longer, way.

Separate u into parallel and perpendicular parts to v :

$$u = \text{proj}_v u + (u - \text{proj}_v u) \quad (2.2)$$

$$= \frac{\langle u, v \rangle}{\langle v, v \rangle} v + (u - \text{proj}_v u) \quad (2.3)$$

Apply the Pythagorean Theorem; reduce fractions as needed.

$$\|u\|^2 = \left\| \frac{\langle u, v \rangle}{\langle v, v \rangle} v \right\|^2 + \|u - \text{proj}_v u\|^2 \quad (2.4)$$

Throw away the right summand.

$$\geq \left\| \frac{\langle u, v \rangle}{\langle v, v \rangle} v \right\|^2 \quad (2.5)$$

The desired result follows from multiplying through by $\langle v, v \rangle$. \square

The following result generalizes the fact that two sides of a triangle sum to longer than the third.

Theorem 8 (Triangle Inequality). Let $u, v \in V$. Then $\|u + v\| \leq \|u\| + \|v\|$.

Proof. Work from the square of the left side.

$$\|u + v\|^2 = \langle u + v, u + v \rangle \quad (2.6)$$

$$= \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle \quad (2.7)$$

$$= \|u\|^2 + \|v\|^2 + \langle u, v \rangle + \langle v, u \rangle \quad (2.8)$$

²That is, the full version. There are cute proofs that just work for \mathbb{R}^n .

Take absolute values of both sides, using the fact that every number is less than or equal to its absolute value, or the scalar “Triangle Inequality.”

$$\|u + v\|^2 \leq \|u\|^2 + \|v\|^2 + |\langle u, v \rangle| + |\langle v, u \rangle| \quad (2.9)$$

Use Cauchy-Schwarz on the inner products.

$$\|u + v\|^2 \leq \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| \quad (2.10)$$

The Triangle Inequality follows after taking square roots. \square

Example of some orthonormal vectors

The following vectors in are orthonormal in \mathbb{C}^3 with the standard inner product:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

2.4 Orthonormal bases and orthogonal transformations

When dealing with inner product spaces, the best basis is an orthonormal basis. Computing coordinates no longer requires solving a system of equations; it is as direct as computing inner products.

Theorem 9 (Computing coordinates in an orthonormal basis). *If $\{v_1, v_2, \dots, v_n\}$ is an orthonormal basis for V , then the coordinates of $v \in V$ relative to this basis are $(\langle v, v_1 \rangle, \langle v, v_2 \rangle, \dots, \langle v, v_n \rangle)$.*

Proof. We will presume a unique decomposition and solve for a single coordinate.

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n \quad (2.11)$$

Take inner products of both sides with v_i on the right.

$$\langle v, v_i \rangle = \langle \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n, v_i \rangle \quad (2.12)$$

Use linearity in the first argument.

$$\langle v, v_i \rangle = \alpha_1 \langle v_1, v_i \rangle + \alpha_2 \langle v_2, v_i \rangle + \dots + \alpha_n \langle v_n, v_i \rangle = \alpha_i \quad (2.13)$$

\square

The following definition generalizes rigid transformations in Euclidean geometry.

Definition 13 (Orthogonal transformation). *Let $f : U \rightarrow V$ be a linear map of inner product spaces. f is called an **orthogonal transformation** if it preserves length: $\|f(u)\| = \|u\|$. If $U = V$, f is called an **orthogonal operator**.*

The following result generalizes the “side-side-side” triangle congruence of classical synthetic geometry—if two triangles have the same side lengths, they have the same angles too.

Theorem 10 (Orthogonal transformation preserves inner product). *Let $f : U \rightarrow V$ be an orthogonal transformation and let $x, y \in U$. Then $\langle f(x), f(y) \rangle = \langle x, y \rangle$.*

Proof. The plan is to realize the desired product as cross terms of a product of sums.

$$\|f(x) + f(y)\|^2 = \|x + y\|^2 \quad (2.14)$$

$$\|f(x)\|^2 + \langle f(x), f(y) \rangle + \langle f(y), f(x) \rangle + \|f(y)\|^2 = \|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2 \quad (2.15)$$

$$\langle f(x), f(y) \rangle + \langle f(y), f(x) \rangle = \langle x, y \rangle + \langle y, x \rangle \quad (2.16)$$

$$\operatorname{Re} \langle f(x), f(y) \rangle = \operatorname{Re} \langle x, y \rangle \quad (2.17)$$

Repeating with the identity $\|f(x) - f(y)\|^2 = \|x - y\|^2$ yields $\operatorname{Im} \langle f(x), f(y) \rangle = \operatorname{Im} \langle x, y \rangle$. \square

This implies that the image of an orthonormal basis is orthonormal, so orthogonal transformations are one-to-one.

Theorem 11. *Let $f : U \rightarrow V$ be an orthogonal transformation, and let $\{u_1, u_2, \dots, u_n\}$ and $\{v_1, v_2, \dots, v_m\}$ be orthonormal bases for U and V , respectively. The matrix of f relative to these bases has orthonormal columns.*

The proof involves a lot of inner product chasing and is in my opinion more technical than illuminating.

Definition 14 (Unitary matrix). *A unitary matrix is a square matrix $U \in \mathbb{C}^n$ with the property that $U^*U = I$.*

The matrix of an orthogonal operator in an orthonormal basis is unitary.

Theorem 12. *Let λ be an eigenvalue of an orthogonal operator $f : V \rightarrow V$. Then $|\lambda| = 1$.*

Proof. Let $f(v) = \lambda v$. Then

$$\langle v, v \rangle = \langle f(v), f(v) \rangle = \langle \lambda v, \lambda v \rangle \quad (2.18)$$

$$= \lambda \bar{\lambda} \langle v, v \rangle, \quad (2.19)$$

so $\lambda \bar{\lambda} = 1$. \square