Lecture 19: Inner Product Spaces

March 28, 2025

Inner Products

In $\mathbb{R}^2$, the scalar (or dot) product of the vectors $x = \binom{x_1}{x_2}$ and $y = \binom{y_1}{y_2}$ is given by

$$x \cdot y = x_1 y_1 + x_2 y_2 = x^T y = y^T x = y \cdot x.$$

We can interpret the *length* $||x||$ of the vector x as the length of the directed line segment from the origin to (x_1, x_2), which by the Theorem of Pythagoras is $\sqrt{x_1^2 + x_2^2}$ or $\sqrt{x \cdot x}$.

Once we have a concept of length of a vector, we can define the *distance* $d(x, y)$ between two vectors x and y as the length of their difference: $d(x, y) = ||x - y||$.

Inner Products

In $\mathbb{R}^2$, the scalar (or dot) product of the vectors $x = \binom{x_1}{x_2}$ and $y = \binom{y_1}{y_2}$ is given by

$$x \cdot y = x_1 y_1 + x_2 y_2 = x^T y = y^T x = y \cdot x.$$

Similarly, from the Cosine Rule we can observe that $x \cdot y = ||x|| \, ||y|| \cos\theta$, where θ is the angle between the directed line segments representing x and y. In particular, x is orthogonal to y (or $x \perp y$) if and only if $x \cdot y = 0$.

So the scalar product encodes geometric information in $\mathbb{R}^2$, and it also provides a mechanism for defining concepts of length, distance and orthogonality on real vector spaces that do not necessarily have an obvious geometric structure.
The scalar product is an example of an inner product.

Real Inner Products

An inner product on a vector space V is a function from $V \times V$ to $\mathbb{R}$ that assigns an element of $\mathbb{R}$ to every ordered pair of elements of V, and has the following properties.

1. Symmetry: $\langle x, y\rangle = \langle y, x\rangle$ for all $x, y \in V$
2. Linearity in both slots (bilinearity): For all $x, y, z \in V$ and all $a, b \in \mathbb{R}$, we have $\langle ax + by, z\rangle = a\langle x, z\rangle + b\langle y, z\rangle$ and $\langle x, ay + bz\rangle = a\langle x, y\rangle + b\langle x, z\rangle$.
3. Non-negativity: $\langle x, x\rangle \geq 0$ for all $x \in V$, and $\langle x, x\rangle = 0$ only if $x = 0_V$.

The ordinary scalar product on $\mathbb{R}^n$ is the best known example of an inner product.

Examples of inner products

1. The ordinary scalar product on $\mathbb{R}^n$.
2. Let C be the vector space of all continuous real-valued functions on the interval $[0, 1]$. The analogue of the ordinary scalar product on C is the inner product given by
$$\langle f, g\rangle = \int_0^1 f(x)g(x)\,dx, \text{ for } f, g \in C.$$
3. On the space $M_{m\times n}(\mathbb{R})$, the *Frobenius inner product* or *trace inner product* is defined by $\langle A, B\rangle = \text{trace}(A^T B)$. Note that trace$A^T B$ is the sum over all positions (i, j) of the products $A_{ij}B_{ij}$. So this is closely related to the ordinary scalar product, if the matrices A and B were regarded as vectors with mn entries over $\mathbb{R}$.

It is possible for a single vector space to have many different inner products defined on it, and if there is any risk of ambiguity we need to specify which one we are considering.

Length, Distance and Scaling in an Inner Product Space

Definition We define the length or norm of any vector v by

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

and we define the distance between the vectors u and v by

$$d(u, v) = ||u - v||.$$

Scaling Every vector v and scalar c satisfy $\boxed{||cv|| = |c|\,||v||}$, since

$$||cv|| = \sqrt{\langle cv, cv\rangle} = \sqrt{c^2\langle v, v\rangle} = |c|\,||v||.$$

So we can adjust the norm of any element of V, while preserving its direction, by multiplying it by a positive scalar.

Definition If v is a non-zero vector in an inner product space V, then

$$\hat{v} := \frac{1}{||v||}v$$

is a unit vector in the direction of v, called the *normalization* of v.

Orthogonality in an inner product space

Let V be a vector space with an inner product $\langle \cdot, \cdot \rangle$ (such as the ordinary scalar product).

Definition We say that the vectors u and v are orthogonal (with respect to $\langle \cdot, \cdot \rangle$) if $\langle u, v \rangle = 0$.

All these definitions are consistent with "typical" geometrically motivated concepts of distance and orthogonality.

Examples

1. $(2, 5)$ and $(5, -2)$ are orthogonal with respect to the ordinary scalar product in $\mathbb{R}^2$.
2. $\sin \pi x$ and $\cos \pi x$ are orthogonal with respect to the scalar product on the space of continuous functions on $[0, 1]$ defined in Lecture 18; this is saying that

$$\int_0^1 sin(\pi x) \cos(\pi x)\, dx = 0 \ \left(= \frac{1}{2\pi} \sin^2(\pi x) \Big|_0^1 \right).$$

Orthogonal Projection

Lemma Let u and v be non-zero vectors in an inner product space V. Then it is possible to write (in a unique way) $v = au + v'$, where a is scalar and v' is orthogonal to u.

- If v is orthogonal to u, take $a = 0$ and $v' = v$.
- If v is a scalar multiple of u, take $au = v$ and $v' = 0$.
- Otherwise, to solve for a and v' in the equation $v = au + v'$ (with $u \perp v'$), take the inner product with u on both sides. Then

$$\langle u, v\rangle = a\langle u, u\rangle + 0 \Longrightarrow a = \frac{\langle u, v\rangle}{||u||^2},\ \ v' = v - \frac{\langle u, v\rangle}{||u||^2}u.$$

We can verify directly that the two components in this expression are orthogonal to each other.

Example Write $u = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $v = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$. Then $u = \begin{bmatrix} \frac{3}{2} \\ -\frac{1}{2} \end{bmatrix} + \begin{bmatrix} \frac{1}{2} \\ \frac{3}{2} \end{bmatrix}$.

Orthogonal projection of one vector on another

Definition

For non-zero vectors u and v in an inner product space V, the vector $\frac{\langle u, v\rangle}{||u||^2}u$ *is called the projection of v on the 1-dimensional space spanned by u. It is denoted by* $\mathrm{proj}_u(v)$ *and it has the property that* $v - \mathrm{proj}_u(v)$ *is orthogonal to u.*

Lemma

$\mathrm{proj}_u(v)$ *is the unique element of $\langle u\rangle$ whose distance from v is minimal.*

Proof Let au be a scalar multiple of u. Then

$$d(au, v)^2 = \langle au - v, au - v\rangle = a^2\langle u, u\rangle - 2a\langle u, v\rangle + \langle v, v\rangle$$

Regarded as a quadratic function of a, this has a minimum when its derivative is 0, i.e. when $2a\langle u, u\rangle - 2\langle u, v\rangle = 0$, when $a = \frac{\langle u, v\rangle}{||u||^2}$.

Orthogonal Bases (the Gram-Schmidt process)

Every finite-dimensional inner product space has an orthogonal basis[1]

We can start with any basis $\{b_1, \dots, b_n\}$, and adjust the elements one by one (by subtracting off orthogonal projections of later vectors on earlier ones). The process ends with an orthogonal basis $\{v_1, \dots, v_n\}$.

1. Set $v_1 = b_1$, and $v_2 = b_2 - \text{proj}_{v_1}(b_2) = b_2 - \frac{\langle v_1, b_2 \rangle}{\langle v_1, v_1 \rangle} v_1$.
   Then the pairs b_1, b_2 and v_1, v_2 span the same space, and $v_1 \perp v_2$.
2. Write $v_3 = b_3 - \text{proj}_{v_1}(b_3) - \text{proj}_{v_2}(b_3)$.
   Then $\{v_1, v_2, v_3\}$ and $\{b_1, b_2, b_3\}$ span the same space, and $v_3 \perp v_1$ and $v_3 \perp v_2$. To see this look at $\langle v_3, v_1 \rangle$ and $\langle v_3, v_2 \rangle$, noting that
   $$v_3 = b_3 - \frac{\langle b_3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle b_3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2.$$
3. Continue: at the kth step, form v_k by subtracting from b_k its projections on $v_1, \dots, v_n$.

[1]This means a basis whose elements are all orthogonal to each other.

Orthogonal projection on a subspace

The result of this process is a basis $\{v_1, \dots, v_n\}$ whose elements satisfy

$$\langle v_i, v_j \rangle = 0 \text{ for } i \neq j$$

We can adjust this basis to a orthonormal basis (consisting of orthogonal unit vectors) by replacing each v_i with its normalization $\hat{v}_i$.
From the Gram-Schmidt process, we have

Theorem

If V is a finite-dimensional inner product space, then V has an orthogonal (or orthonormal) basis.

Now let W be a subspace of V, and let $v \in V$. The orthogonal projection of v on W, denoted $\text{proj}_W(v)$, is defined to be the unique element u of W for which

$$v = u + v',$$

and $v' \perp w$ for all $w \in W$.

Calculating the projection on a subspace

Example In $\mathbb{R}^3$, find the unique point of the plane $W : x + 2y - z = 0$ that is nearest to the point $v : (1, 2, 2)$.

Solution First find an orthogonal basis for W: for example $\{b_1, b_2\}$, where

$$b_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \; b_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

Then

$$\begin{aligned} \text{proj}_W(v) &= \text{proj}_{b_1}(v) + \text{proj}_{b_2}(v) \\ &= \frac{\langle b_1, v \rangle}{\langle b_1, b_1 \rangle} b_1 + \frac{\langle b_2, v \rangle}{\langle b_2, b_2 \rangle} b_2 \\ &= \frac{3}{2} b_1 - \frac{3}{3} b_2 \\ &= \left(\frac{3}{2}, 0, \frac{3}{2} \right) - (1, -1, -1) = \left(\frac{1}{2}, 1, \frac{5}{2} \right). \end{aligned}$$

$\text{proj}_W(v)$ is the nearest point of W to v

Let $u = \text{proj}_W(v)$ and let w be any element of W.
Note that $v - u$ is orthogonal to both w and u, hence to $w - u$. Then

$$\begin{aligned} d(v,w)^2 &= \langle v-w, v-w\rangle \\ &= \langle (v-u)+(u-w), (v-u)+(u-w)\rangle \\ &= \langle v-u, v-u\rangle + \cancel{2\langle v-u, u-w\rangle} + \langle u-w, u-w\rangle \\ &= \langle v-u, v-u\rangle + \langle u-w, u-w\rangle \\ &\geq d(v,u)^2, \end{aligned}$$

with equality only if $w = u = \text{proj}_W(v)$.