Let V be a vector space over $\mathbb{R}$. An inner product on 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. We write $\langle x, y\rangle$ for the inner product of x and y, and write the function as $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$.

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 $x = 0_V$.

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$ATB$ 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 Orthogonality

Given a real vector space and equipped with an inner product $\langle \cdot, \cdot, \rangle$, we make the following two definitions.

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

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

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

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

Unit Vectors and Scaling

An element v of V is referred to as a unit vector if $||v|| = 1$.
The norm of elements of V has the property that $||cv|| = |c|z, ||v||$ for any vector v and real scalar c. To see this we can note that

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

So we can adjust the norm of any element of V, while preserving its direction, by multplying 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 same direction as v, referred to as the *normalization* of v.

Orthogonal Projection

Lemma Let u and v be non-zero vectors in an inner product space V. Then it is possible to write $v = au + v'$, 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 the scalar a) in the equation $v = au + v'$ (with $u \perp v'$), take the inner product with u on both sides. Then

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

We conclude that $v = \frac{\langle u,v\rangle}{||u||^2}u + (v - \frac{\langle u,v\rangle}{||u||^2}u)$. We can verify directly that the two components in this expression are orthogonal to each other.

Example In $\mathbb{R}^2$, write $u = \binom{2}{1}$ and $v = \binom{6}{-2}$.

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 $\text{proj}_u(v)$ and it has the property that $v - \text{proj}_u(v)$ is orthogonal to u.

Lemma

$\text{proj}_u(v)$ *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)

Let V be a finite-dimensional inner product space, with a given basis $\mathcal{B} = \{b_1, b_2, \dots b_n\}$.

A basis $\mathcal{B}$ is called orthogonal if its elements are all orthogonal to each other.

We can adjust $\mathcal{B}$ to an orthogonal basis $\mathcal{B}' = \{v_1, \dots, v_n\}$ as follows.

1. Write $v_1 = b_1$.
2. Write $v_2 = b_2 - \text{proj}_{v_1}(v_2) = b_2 - \frac{\langle b_1, b_2 \rangle}{||b_1||^2} b_1$.

   Then the pairs b_1, b_2 and v_1, v_2 span the same space, and $v_1 \perp v_2$.
3. Write $v_3 = b_3 - \text{proj}_{v_1}(b_3) - \text{proj}_{v_2}(b_3)$.

   Then the sets v_1, v_2, v_3 and b_1, b_2, b_3 span the same space, and v_3 is orthogonal to both v_1 and v_2.

   To see this note that
   $$\langle v_1, v_3 \rangle = \langle v_1, b_3 \rangle - \frac{\langle v_1, b_3 \rangle}{\langle v_1, v_1 \rangle} \langle v_1, v_1 \rangle - c\cancel{\langle v_1, v_2 \rangle}$$
4. Continue in this way - at the kth step, form v_k by subtracting from b_k its projections on each of $v_1, \dots, v_n$.

Orthogonal projection on a subspace

The result of this process is a basis $\{v_1, \ldots, v_n\}$ 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-Scmidt 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 $w \in W$.