# Chapter V : Orthogonality

The prototypical example of a vector space is $\mathbb{R}^n$. By extracting some of its features, we were led to define abstract vector spaces. Essentially, we were studying abstract and higher-dimensional analogues of lines & planes.

However, $\mathbb{R}^n$ has much more structure than just a vector space. In particular, one can defines lengths and angles. The idea is to equip a vector space w/ an "inner product" $V \times V \longrightarrow F$ – a bilinear form with certain properties. The definition will require $F$ to contain an "ordered" subfield. For this reason, we will only consider $F = \mathbb{R}$ or $F = \mathbb{C}$. So for the rest of this chapter, every vector space will be real or complex.

In $\mathbb{R}^2$, the usual inner product is the dot product:

$$\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2$$

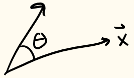
One defines the length (or the norm) of a vector in terms of the dot product:

$$\|\vec{x}\| := \sqrt{\vec{x} \cdot \vec{x}}$$

Then one can define the angle between two vectors

$\vec{y}$

$$\cos\theta = \frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\,\|\vec{y}\|}$$



The point is that lengths and angles can be defined using the dot product. We will generalize the dot product to the notion of an "inner product" on an abstract real or complex vector space.

## 5.1 Inner product Spaces

**5.1.1 Definition** An inner product on a real vector space $V$ is a bilinear form

$$\langle -,- \rangle : V \times V \longrightarrow \mathbb{R}$$

satisfying

(1) $\langle x,y\rangle = \langle y,x\rangle \quad \forall x,y \in V$

(2) $\langle x,x\rangle \geq 0 \quad \forall x \in V$

(3) $\langle x,x\rangle = 0$ if and only if $x=0$.

A real inner product space is a real vector space equipped with an inner product. $\square$

Remark: A bilinear form satisfying (1) is called symmetric. Condition (2) says it is positive-definite. Notice that condition (2) requires a notion of order. This is one reason why we are working over $\mathbb{R}$. Notice that a symmetric bilinear form is automatically reflexive (see Weekly 3 Def 2). Condition (3) implies that

$\mathrm{rad}(V) = \{0\}$, so a (real) inner product is in particular a nondegenerate bilinear form.

**5.1.2 Definition** An inner product on a complex vector space $V$ is a function

$$\langle -,- \rangle : V\times V \longrightarrow \mathbb{C}$$

Satisfying

(0) $\langle -,- \rangle$ is linear in the first argument:

$$\langle \lambda x+\mu y, z\rangle = \lambda\langle x,z\rangle + \mu\langle y,z\rangle \quad \forall x,y,z\in V,\ \forall \lambda,\mu\in\mathbb{C}.$$

(1) $\langle x,y\rangle = \overline{\langle y,x\rangle} \quad \forall x,y\in V$

(2) $\langle x,x\rangle \geq 0 \quad \forall x\in V$

(3) $\langle x,x\rangle = 0$ if and only if $x=0$.

complex conjugation

(1) implies $\langle x,x\rangle\in\mathbb{R}$ so 2 makes sense.

A complex inner product space is a complex vector space equipped with an inner product. □

**Remark** (0) & (1) imply that $\langle -,- \rangle$ is "conjugate linear" in the second argument:

$$\langle x, \lambda y+\mu z\rangle \overset{(1)}{=} \overline{\langle \lambda y+\mu z, x\rangle} \overset{(0)}{=} \bar\lambda\,\overline{\langle y,x\rangle} + \bar\mu\,\overline{\langle z,x\rangle} \overset{(1)}{=} \bar\lambda\langle x,y\rangle+\bar\mu\langle x,z\rangle.$$

A map $\langle -,- \rangle : V\times V \longrightarrow \mathbb{C}$ that is linear in the first arg and conjugate linear in the second is called a sesquilinear form.

**Remark** By (1), $\langle x,x\rangle = \overline{\langle x,x\rangle}$ for any $x \in V$. So $\langle x,x\rangle \in \mathbb{R}$ and "$\geq$" in (2) makes sense!

**Terminology** A real inner product space is also called a Euclidean space. A complex inner product space is also called a Hermitian space. If I just say "inner product space" it could be real or complex and I will use $\mathbb{F}$ to denote either $\mathbb{R}$ or $\mathbb{C}$.

**5.1.3 Definition** Let $V$ be an inner product space. We define a function $\|\cdot\| : V \longrightarrow \mathbb{R}$ via

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

for any $x \in V$. We call $\|\cdot\|$ the **norm** on $V$ induced by $\langle -,-\rangle$. $\square$

**Remark** A norm on a real or complex vector space $V$ is a function $\|\cdot\| : V \longrightarrow \mathbb{R}$ satisfying

(1) $\|x+y\| \leq \|x\| + \|y\| \quad \forall x, y \in V$

(2) $\|\alpha x\| = |\alpha| \|x\| \quad \forall x \in V, \quad \alpha \in \mathbb{F}$

(3) $\|x\| \geq 0$ and $\|x| = 0$ iff $x = 0$.

One can check that the norm induced by an inner product is actually a norm in the above sense. A norm gives an appropriate notion of the length of a vector.

**5.1.4 Example**

- $V = \mathbb{R}^n$, $\langle \vec{x}, \vec{y} \rangle := \sum_{i=1}^{n} x_i y_i$ "standard inner product on $\mathbb{R}^n$"
- $V = \mathbb{C}^n$, $\langle \vec{z}, \vec{w} \rangle := \sum_{i=1}^{n} z_i \overline{w_i}$ "Standard inner product on $\mathbb{C}^n$"
- $V = C([0,1], \mathbb{C})$ the complex vector space of continuous functions $[0,1] \to \mathbb{C}$

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

## 5.2 Orthogonality

**5.2.1 Definition** Let $V$ be an inner product space. Two vectors $x, y \in V$ are called orthogonal if

$$\langle x, y \rangle = 0.$$

This condition is symmetric: $\langle x, y \rangle = 0$ iff $\langle y, x \rangle = 0$. $\square$

**5.2.2 Definition** A subset $S \subseteq V$ in an inner product space $V$ is called orthogonal if $\langle x, y \rangle = 0$ whenever $x, y \in S$ with $x \neq y$ and orthonormal if

$$\langle x, y \rangle = \begin{cases} 1, & x = y \\ 0, & x \neq y. \end{cases}$$

for all $x, y \in S$. $\square$

**Note:** $\langle x, y \rangle = 1$ iff $\|x\| = 1$, so an orthonormal set of vectors is an orthogonal set of unit vectors - vectors of norm equal to 1.

**Remark:**

- An orthogonal set $S$ is linearly independent. If $\sum_{i=1}^{n} \alpha_i x_i = 0$, then

$$
\begin{aligned}
0 &= \langle \textstyle\sum \alpha_i x_i , x_j \rangle \\
&= \sum_{i=1}^{n} \alpha_i \langle x_i , x_j \rangle \\
&= \alpha_j \langle x_j , x_j \rangle .
\end{aligned}
$$

Since $S$ is indep, $x_j \neq 0$, so $\langle x_j , x_j \rangle > 0$. Thus, $\alpha_j = 0$.

- The zero vector is orthogonal to every vector.
- If $x \neq 0$, the set $\{x\}$ is orthogonal and the set $\{\frac{x}{\|x\|}\}$ is orthonormal.

**5.2.3 Definition** An orthonormal basis is an orthonormal subset $B \subseteq V$ which is a basis for $V$. Equivalently, an orthonormal basis is an orthonormal spanning set. □

**5.2.4 Theorem** Every finite-dimensional inner product space has an orthonormal basis.

Proof. Let $\{x_1, \dots, x_n\}$ be a basis for $V$. We apply the "Gram-Schmidt Algorithm" in order to produce an orthonormal basis $\{u_1, \dots, u_n\}$. We will construct, by induction on $n$, an orthonormal set $\{u_1, \dots, u_n\}$ such that

$$\operatorname{span}\{u_1, \dots, u_n\} = \operatorname{span}\{v_1, \dots, v_n\}$$

If $n=1$, then $u_1 := \frac{x_1}{\|x_1\|}$. Suppose $n \geq 2$. Let

$$y_n = x_n - \sum_{i=1}^{n-1} \langle x_n , u_i \rangle u_i$$

an set $u_n = \frac{y_n}{\|y_n\|}$. By inductive hypothesis, $\{u_1,\dots,u_{n-1}\}$ is orthonormal with $\operatorname{span}\{u_1,\dots,u_{n-1}\} = \operatorname{span}\{v_1,\dots,v_{n-1}\}$. Since $\|u_n\|=1$, we just need to show that $\langle u_n,u_i\rangle = 0 \ \forall i=1,\dots,n-1$. We have

$$\langle u_n, u_i\rangle = \frac{1}{\|y_n\|}\langle y_n,u_i\rangle$$

and

$$\begin{aligned}\langle y_n,u_i\rangle &= \langle x_n,u_i\rangle - \sum_{j=1}^{n-1}\langle x_n,u_j\rangle\langle u_i,u_j\rangle \\ &= \langle x_n,u_i\rangle - \langle x_n,u_i\rangle\langle u_i,u_i\rangle \\ &= \langle x_n,u_i\rangle - \langle x_n,u_i\rangle\cdot 1 \\ &= 0.\end{aligned}$$

Thus, $\{u_1,\dots,u_n\}$ is an orthonormal set. Moreover, $x_n \in \operatorname{span}\{u_1,\dots,u_n\}$ by construction of $u_n$. Thus,

$$\operatorname{span}\{x_1,\dots,x_n\} \overset{\text{inductive hypothesis}}{=} \operatorname{Span}\{u_1,\dots,u_{n-1},x_n\} \subseteq \operatorname{Span}\{u_1,\dots,u_n\}. \qquad \square$$

**Remark** Any subspace $W$ of an inner product is an inner product space by restricting $\langle -,-\rangle$ to $W\times W$. Thus, any subspace of an inner product space has an orthonormal basis.

**Remark** Suppose $B=\{u_1,\dots,u_n\}$ is an orthonormal basis of $V$. Let $x\in V$ and suppose $[x]_B = (\alpha_1,\dots,\alpha_n)$. Then

$$\langle x,u_i\rangle = \sum_{j=1}^{n}\alpha_i\langle u_j,u_i\rangle = \alpha_i.$$

Thus, $[x]_B = (\langle x,u_1\rangle, \dots, \langle x,u_n\rangle)$, that is,

$$\boxed{x = \sum_{i=1}^{n} \langle x, u_i \rangle u_i .}$$

In other words, the linear functionals $\langle -, u_i \rangle : V \to \mathbb{F}$ are the vectors in the dual basis $\mathcal{B}^*$ for $V^*$.

We will pretty much exclusively use orthonormal bases from now on, since they exist and they are the most natural and convenient to use.

**5.2.5 Definition** If $V$ is an inner product space and $S \subseteq V$, then we define

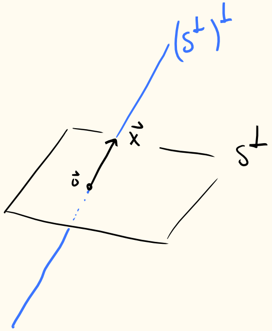
$$S^{\perp} := \{ v \in V \mid \langle x, v \rangle = 0 \;\; \forall x \in S \} .$$

If $W$ is a subspace of $V$, then $W^{\perp}$ is called the orthogonal complement of $W$. $\square$

**Facts:**
- $S^{\perp}$ is a subspace of $V$, even if $S$ is not.
- $S \subseteq (S^{\perp})^{\perp} \Rightarrow \operatorname{span}(S) \subseteq (S^{\perp})^{\perp}$

**5.2.6 Example**

$S = \{\vec{x}\}$



- $\{0\}^{\perp} = V$
- $V^{\perp} = \{0\}$

**5.2.7 Theorem** If $W$ is a subspace of a finite-dimensional inner product space $V$, then $W^\perp$ is a complementary subspace of $W$

$$V = W \oplus W^\perp$$

and $W = (W^\perp)^\perp$.

Proof. Let $x \in W \cap W^\perp$. Then $x = 0$ since $\langle x, x\rangle = 0$. To show that $W + W^\perp = V$, choose an orthonormal basis $u_1, \dots, u_k$ for $W$ and extend to an orthonormal basis

$$u_1, \dots, u_k, y_1, \dots, y_m$$

for $V$. It suffices to show that $y_i \in W^\perp$ for all $i = 1, \dots, m$. Let $x \in W$. Then

$$\begin{aligned} \langle x, y_i\rangle &= \left\langle \sum_{j=1}^{k} \langle x, u_j\rangle u_j, y_i \right\rangle \\ &= \sum_{j=1}^{k} \langle x, u_j\rangle \langle u_j, y_i\rangle \\ &= 0 \qquad (\text{since } u_j \in W,\ y_i \in W^\perp). \end{aligned}$$

Finally, $(W^\perp)^\perp \subseteq W$. Let $x \in (W^\perp)^\perp$. Write

$$x = \sum_{i=1}^{k} \langle x, u_i\rangle u_i + \sum_{i=1}^{m} \langle x, y_j\rangle y_j$$

Then

$$\begin{aligned} 0 = \langle x, y_j\rangle &= \left\langle \sum_i \langle x, u_i\rangle u_i + \sum_i \langle x, y_i\rangle y_i, y_j \right\rangle \\ &= \sum_i \langle x, u_i\rangle \langle u_i, y_j\rangle + \sum_i \langle x, y_i\rangle \langle y_i, y_j\rangle \\ &= \langle x, y_j\rangle. \end{aligned}$$

Thus, $x = \sum_{i=1}^{k} \langle x, u_i\rangle u_i \in W$. $\square$

**Remark:** Thus, in a finite-dimensional inner product space, every subspace $W \subseteq V$ has a canonical choice of complementary subspace, namely $W^\perp$.

**5.2.8 Definition** A projection on a vector space $V$ is a linear operator $P : V \to V$ satisfying $P^2 = P$. $\square$

**5.2.9. Proposition** If $P$ is a projection, then

$$V = \operatorname{im}(P) \oplus \ker(P).$$

Conversely, given $V = W \oplus W'$, there is a unique projection $P : V \to V$ satisfying

$$\operatorname{im}(P) = W \quad \& \quad \ker(P) = W'.$$

It is called the projection onto $W$ along $W'$. $\square$

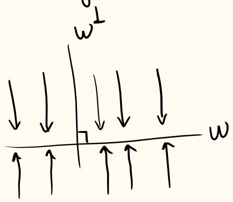
In other words, a projection is completely determined by its image and kernel, which can be arbitrary complementary subspaces.

**5.2.10 Definition** If $V$ is an inner product space, then any subspace $W$ has a complement $W^\perp$. The projection onto $W$ along $W^\perp$ is called the orthogonal projection onto $W$ and is denoted by

$$P_W : V \longrightarrow V$$

$\square$

**Remark** Since $W^\perp$ is a canonical choice for a complement of $W$, an orthogonal projection is completely determined by its image.



Explicitly, if $u_1, \dots, u_n$ is an orthonormal basis for $W$, then

$$P_W(v) = \sum_{i=1}^{n} \langle v, u_i \rangle u_i .$$

Let $P_i : V \to V$ be projection onto $\operatorname{span}\{u_i\}$ where $u_1, \dots, u_n$ is an orthonormal basis for $W$. Then

$$P_i(v) = \langle v, u_i \rangle u_i .$$

Thus, $P_W = \sum_{i=1}^{n} P_i$. More generally, if $W = W_1 \oplus \cdots \oplus W_k$ are pairwise orthogonal subspaces, then $P_W = \sum_{i=1}^{k} P_{W_i}$.

## 5.3 Linear Functionals on Inner Product Spaces

Let $V$ be an inner product space. For any $y \in V$, the map $\langle -, y \rangle : V \to \mathbb{F}$ is a linear function. This defines a function

$$\begin{aligned} V &\longrightarrow V^* \\ y &\longmapsto \langle -, y \rangle . \end{aligned}$$

### 5.3.1 Theorem

If $V$ is a finite-dimensional inner product space, then

$$\begin{aligned} \Phi : V &\longrightarrow V^* \\ y &\longmapsto \langle -, y \rangle \end{aligned}$$

is a bijection.

Proof. The proof is similar to Weekly 3.2. □

**Remark** $\Phi(y_1+y_2) = \langle -, y_1+y_2\rangle = \langle -, y_1\rangle + \langle -, y_2\rangle = \Phi(y_1)+\Phi(y_2)$

$$\Phi(\lambda y) = \langle -, \lambda y\rangle = \bar{\lambda}\langle -, y\rangle = \bar{\lambda}\,\Phi(y).$$

So when $\mathbb{F} = \mathbb{C}$, $\Phi$ is not a linear map - it is "conjugate linear". When $\mathbb{F} = \mathbb{R}$, $\Phi$ is a linear map and is therefore a vector space isomorphism. You proved this in Weekly 3.2 since $\langle -,-\rangle$ is a nondegenerate reflexive bilinear form.

**Remark** We can use $\Phi$ to equip $V^*$ with an inner product $\langle -,-\rangle : V^* \times V^* \longrightarrow \mathbb{F}$. Let $f, g \in V^*$. There exist unique $y_1, y_2 \in V$ s.t. $f = \Phi(y_1)$ and $g = \Phi(y_2)$. Then

$$\langle f, g\rangle := \overline{\langle y_1, y_2\rangle} = \langle y_2, y_1\rangle.$$

This definition makes $V^*$ into an inner product space.

Significance of the Theorem We can use $V \xrightarrow{\sim} V^*$ to turn linear functionals into vectors.

5.3.2 Example Let $W \subseteq V$ be a subspace.

$$W^0 := \{ f \in V^* \mid f(W) = \{0\} \}$$ is the annihilator of $W$. Under the bijection

$$\begin{tikzcd} V \arrow[r, "\sim"] & V^* \\ W^\perp \arrow[r, leftrightarrow] & W^0 \end{tikzcd}$$

the annihilator corresponds to the orthogonal complement. See Weekly 3.

**5.3.3 Example** If $L: V \to V$ is a linear operator, there is a unique linear operator $L^*: V^* \to V^*$.

$$\begin{tikzcd} f \arrow[r, mapsto] & f \circ L \end{tikzcd}$$

Identify $V \xrightarrow{\sim} V^*$ via $\Phi$. Then the composition

$$\begin{tikzcd}
V \arrow[r, "\Phi", "\sim"'] & V^* \arrow[r, "L^*"] & V^* \arrow[r, "\Phi^{-1}", "\sim"'] & V \\
v \arrow[r, mapsto] & \langle -, v \rangle & f \arrow[r, mapsto] & \text{unique } y : f = \langle -, y \rangle . \\
 & f \arrow[r, mapsto] & f \circ L & \\
v \arrow[rrr, mapsto] & & & \langle L(y, v \rangle
\end{tikzcd}$$

is a new linear operator on $V$. We use the same notation $L^*: V \to V$. We call it the adjoint of $L$.

**5.3.4 Theorem** Let $L: V \to V$ be a linear operator on a finite-dimensional inner product space. Then there exists a unique linear operator $L^*: V \to V$ satisfying

$$\langle Lx, y \rangle = \langle x, L^* y \rangle \quad \forall x, y \in V.$$

Explicitly, given $y \in V$, $L^* y$ is the unique vector in $V$ such that $\langle -, L^* y \rangle = \langle T(-), y \rangle$. The operator $L^*$ is called the adjoint of $L$.

Proof. Exercise. $\square$