Inner Product Spaces

Isaiah Lankham, Bruno Nachtergaele, Anne Schilling

(March 2, 2007)

The abstract definition of vector spaces only takes into account algebraic properties for the addition and scalar multiplication of vectors. For vectors in $\mathbb{R}^n$, for example, we also have geometric intuition which involves the length of vectors or angles between vectors. In this section we discuss inner product spaces, which are vector spaces with an inner product defined on them, which allow us to introduce the notion of length (or norm) of vectors and concepts such as orthogonality.

1 Inner product

In this section V is a finite-dimensional, nonzero vector space over $\mathbb{F}$.

Definition 1. An **inner product** on V is a map

$$\begin{aligned} \langle \cdot, \cdot \rangle : V \times V &\to \mathbb{F} \\ (u, v) &\mapsto \langle u, v \rangle \end{aligned}$$

with the following properties:

1. **Linearity in first slot:** $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$ and $\langle au, v\rangle = a\langle u, v\rangle$;
2. **Positivity:** $\langle v, v\rangle \geq 0$ for all $v \in V$;
3. **Positive definiteness:** $\langle v, v\rangle = 0$ if and only if $v = 0$;
4. **Conjugate symmetry:** $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$.

Remark 1. Recall that every real number $x \in \mathbb{R}$ equals its complex conjugate. Hence for real vector spaces the condition about conjugate symmetry becomes symmetry.

Definition 2. An **inner product space** is a vector space over $\mathbb{F}$ together with an inner product $\langle \cdot, \cdot \rangle$.

Example 1. $V = \mathbb{F}^n$

$$u = (u_1, \dots, u_n), v = (v_1, \dots, v_n) \in \mathbb{F}^n$$

Then

$$\langle u, v \rangle = \sum_{i=1}^{n} u_i \overline{v}_i.$$

For $\mathbb{F} = \mathbb{R}$, this is the usual dot product

$$u \cdot v = u_1 v_1 + \cdots + u_n v_n.$$

For a fixed vector $w \in V$, one may define the map $T : V \to \mathbb{F}$ as $Tv = \langle v, w \rangle$. This map is linear by condition 1 of Definition 1. This implies in particular that $\langle 0, w \rangle = 0$ for every $w \in V$. By the conjugate symmetry we also have $\langle w, 0 \rangle = 0$.

Lemma 2. *The inner product is anti-linear in the second slot, that is,* $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$ *for all* $u, v, w \in V$ *and* $\langle u, av \rangle = \overline{a} \langle u, v \rangle$.

Proof. For the additivity note that

$$\begin{aligned} \langle u, v + w \rangle &= \overline{\langle v + w, u \rangle} = \overline{\langle v, u \rangle + \langle w, u \rangle} \\ &= \overline{\langle v, u \rangle} + \overline{\langle w, u \rangle} = \langle u, v \rangle + \langle u, w \rangle. \end{aligned}$$

Similarly,

$$\langle u, av \rangle = \overline{\langle av, u \rangle} = \overline{a \langle v, u \rangle} = \overline{a}\,\overline{\langle v, u \rangle} = \overline{a} \langle u, v \rangle.$$

□

Note that the convention in physics is often different. There the second slot is linear, whereas the first slot is anti-linear.

2 Norms

The norm of a vector is the analogue of the length. It is formally defined as follows.

Definition 3. Let V be a vector space over $\mathbb{F}$. A map

$$\begin{aligned} \| \cdot \| : V &\to \mathbb{R} \\ v &\mapsto \|v\| \end{aligned}$$

is a **norm** on V if

1. $\|v\| = 0$ if and only if $v = 0$;

2. $\|av\| = |a|\|v\|$ for all $a \in \mathbb{F}$ and $v \in V$;
3. **Triangle inequality** $\|v + w\| \leq \|v\| + \|w\|$ for all $v, w \in V$.

Note that in fact $\|v\| \geq 0$ for all $v \in V$ since

$$0 = \|v - v\| \leq \|v\| + \|-v\| = 2\|v\|.$$

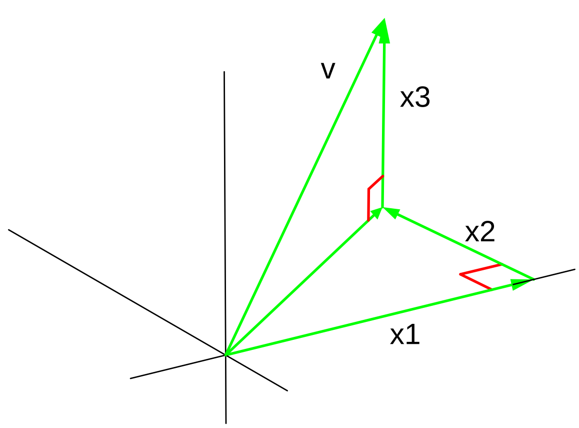
Next we want to show that a norm can in fact be defined from an inner product via

$$\|v\| = \sqrt{\langle v, v\rangle} \quad \text{for all } v \in V.$$

Properties 1 and 2 follow easily from points 1 and 3 of Definition 1. The triangle inequality requires proof (which we give in Theorem 5).

Note that for $V = \mathbb{R}^n$ the norm is related to what you are used to as the distance or length of vectors. Namely for $v = (x_1, \ldots, x_n) \in \mathbb{R}^n$ we have

$$\|v\| = \sqrt{x_1^2 + \cdots + x_n^2}.$$



3 Orthogonality

Using the inner product, we can now define the notion of orthogonality, prove the Pythagorean theorem and the Cauchy-Schwarz inequality which will enable us to prove the triangle inequality. This will show that $\|v\| = \sqrt{\langle v, v\rangle}$ indeed defines a norm.

Definition 4. Two vectors $u, v \in V$ are **orthogonal** ($u \perp v$ in symbols) if and only if $\langle u, v\rangle = 0$.

Note that the zero vector is the only vector that is orthogonal to itself. In fact, the zero vector is orthogonal to all vectors $v \in V$.

Theorem 3 (Pythagorean Theorem)**.** *If $u, v \in V$ with $u \perp v$, then*

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

Proof. Suppose $u, v \in V$ such that $u \perp v$. Then

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

□

Note that the converse of the Pythagorean Theorem holds for real vector spaces, since in this case $\langle u, v\rangle + \langle v, u\rangle = 2\mathrm{Re}\langle u, v\rangle = 0$.

Given two vectors $u, v \in V$ with $v \neq 0$ we can uniquely decompose u as a piece parallel to v and a piece orthogonal to v. This is also called the **orthogonal decomposition**. More precisely

$$u = u_1 + u_2$$

so that $u_1 = av$ and $u_2 \perp v$. Namely write $u_2 = u - u_1 = u - av$. For u_2 to be orthogonal to v we need

$$0 = \langle u - av, v\rangle = \langle u, v\rangle - a\|v\|^2.$$

Solving for a yields $a = \langle u, v\rangle / \|v\|^2$, so that

$$u = \frac{\langle u, v\rangle}{\|v\|^2} v + \left(u - \frac{\langle u, v\rangle}{\|v\|^2} v\right). \tag{1}$$

Theorem 4 (Cauchy-Schwarz inequality)**.** *For all $u, v \in V$ we have*

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

Furthermore, equality holds if and only if u and v are linearly dependent, i.e., are scalar multiples of each other.

Proof. If $v = 0$, then both sides of the inequality are zero. Hence assume that $v \neq 0$. Consider the orthogonal decomposition

$$u = \frac{\langle u, v\rangle}{\|v\|^2} v + w$$

where $w \perp v$. By the Pythagorean theorem we have

$$\|u\|^2 = \left\| \frac{\langle u, v\rangle}{\|v\|^2} v \right\|^2 + \|w\|^2 = \frac{|\langle u, v\rangle|^2}{\|v\|^2} + \|w\|^2 \geq \frac{|\langle u, v\rangle|^2}{\|v\|^2}.$$

Multiplying both sides by $\|v\|^2$ and taking the square root yields the Cauchy-Schwarz inequality.

Note that we get equality in the above arguments if and only if $w = 0$. But by (1) this means that u and v are linearly dependent. □

The Cauchy-Schwarz inequality has many different proofs. Here is another one.

Proof. For given $u, v \in V$ consider the norm square of the vector $u + re^{i\theta}v$,

$$0 \leq \|u + re^{i\theta}v\|^2 = \|u\|^2 + r^2\|v\|^2 + 2\mathrm{Re}(re^{i\theta}\langle u, v\rangle).$$

Since $\langle u, v\rangle$ is a complex number, one can choose θ so that $e^{i\theta}\langle u, v\rangle$ is real. Hence the right hand side is a parabola $ar^2 + br + c$ with real coefficients. It will lie above the real axis, i.e. $ar^2 + br + c \geq 0$, if it does not have any real solutions for r. This is the case when the discriminant satisfies $b^2 - 4ac \leq 0$. In our case this means

$$4|\langle u, v\rangle|^2 - 4\|u\|^2\|v\|^2 \leq 0.$$

Equality only holds if r can be chosen such that $u + re^{i\theta}v = 0$, which means that u and v are scalar multiples. □

We now come to the proof of the triangle inequality which shows that $\|v\| = \sqrt{\langle v, v\rangle}$ indeed defines a norm.

Theorem 5 (Triangle inequality)**.** *For all $u, v \in V$ we have*

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

Proof. By straighforward calculation we obtain

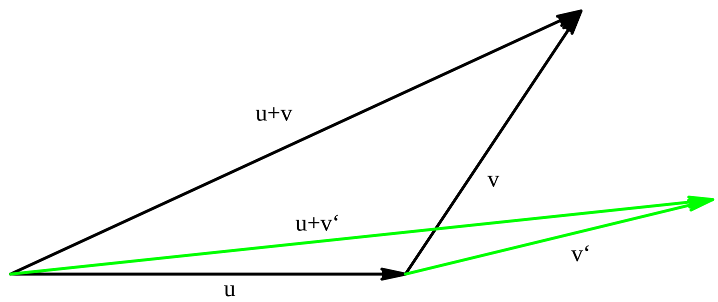
$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v\rangle = \langle u, u\rangle + \langle v, v\rangle + \langle u, v\rangle + \langle v, u\rangle \\ &= \langle u, u\rangle + \langle v, v\rangle + \langle u, v\rangle + \overline{\langle u, v\rangle} = \|u\|^2 + \|v\|^2 + 2\mathrm{Re}\langle u, v\rangle.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

Note that $\mathrm{Re}\langle u, v\rangle \leq |\langle u, v\rangle|$ so that using the Cauchy-Schwarz inequality we obtain

$$\|u + v\|^2 \leq \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| = (\|u\| + \|v\|)^2.$$

Taking the square root of both sides gives the triangle inequality. □

Remark 6. Note that equality holds for the triangle inequality if and only if $v = ru$ or $u = rv$ for $r \geq 0$. Namely, equality in the proof happens only if $\langle u, v \rangle = \|u\|\|v\|$ which is equivalent to the above statement.



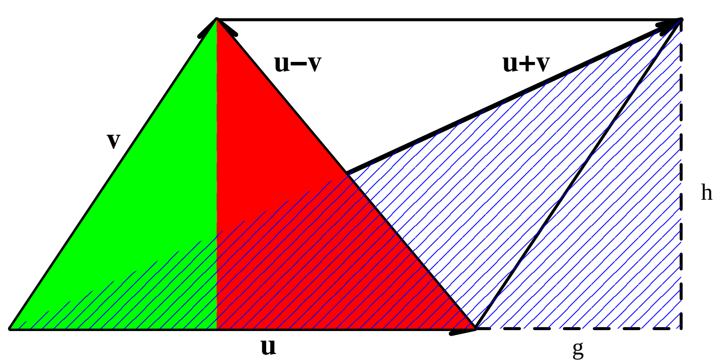
Theorem 7 (Parallelogram equality)**.** *For all $u, v \in V$ we have*

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

Proof. By direct calculation

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

□



4 Orthonormal bases

We now define the notion of orthogonal and orthonormal bases of an inner product space. As we will see later, orthonormal bases have very special properties that simplify many calculations.

Definition 5. Let V be an inner product space with inner product $\langle \cdot, \cdot \rangle$. A list of nonzero vectors $(e_1, \ldots, e_m)$ of V is called **orthogonal** if

$$\langle e_i, e_j \rangle = 0 \quad \text{for all } 1 \leq i \neq j \leq m.$$

The list $(e_1, \ldots, e_m)$ is called **orthonormal** if

$$\langle e_i, e_j \rangle = \delta_{i,j} \quad \text{for all } i, j = 1, \ldots, m,$$

where δ_{ij} is the Kronecker delta symbol and is 1 if $i = j$ and zero otherwise.

Proposition 8. *Every orthogonal list of nonzero vectors in V is linearly independent.*

Proof. Let $(e_1, \ldots, e_m)$ be an orthogonal list of vectors in V and suppose $a_1, \ldots, a_m \in \mathbb{F}$ are such that

$$a_1 e_1 + \cdots + a_m e_m = 0.$$

Then

$$0 = \|a_1 e_1 + \cdots + a_m e_m\|^2 = |a_1|^2 \|e_1\|^2 + \cdots + |a_m|^2 \|e_m\|^2$$

Note that $\|e_k\| > 0$ for all $k = 1, \ldots, m$ since every e_k is a nonzero vector. Also $|a_k|^2 \geq 0$. Hence the only solution to this equation is $a_1 = \cdots = a_m = 0$. □

Definition 6. An **orthonormal basis** of a finite-dimensional inner product space V is an orthonormal list of vectors that is basis (i.e., in particular spans V).

Clearly any orthonormal list of length $\dim V$ is a basis of V.

Example 2. The canonical basis of $\mathbb{F}^n$ is orthonormal.

Example 3. The list $((\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}))$ is an orthonormal basis of $\mathbb{R}^2$.

The next theorem shows that the coefficients of a vector $v \in V$ in terms of an orthonormal basis are easy to compute via the inner product.

Theorem 9. *Let $(e_1, \ldots, e_n)$ be an orthonormal basis of V. Then for all $v \in V$ we have*

$$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n$$

and $\|v\|^2 = \sum_{k=1}^{n} |\langle v, e_k \rangle|^2$.

Proof. Let $v \in V$. Since $(e_1, \ldots, e_n)$ is a basis of V there exist unique scalars $a_1, \ldots, a_n \in \mathbb{F}$ such that

$$v = a_1 e_1 + \cdots + a_n e_n.$$

Taking the inner product on both sides with respect to e_k yields $\langle v, e_k \rangle = a_k$. □

5 The Gram-Schmidt orthogonalization procedure

We now come to a very important algorithm, called the Gram-Schmidt orthogonalization procedure. This algorithm makes it possible to construct for each list of linearly independent vectors (or a basis) a corresponding orthonormal list (or orthonormal basis).

Theorem 10. *If $(v_1, \ldots, v_m)$ is a linearly independent list of vectors in V, then there exists an orthonormal list $(e_1, \ldots, e_m)$ such that*

$$\textit{span}(v_1, \ldots, v_k) = \textit{span}(e_1, \ldots, e_k) \quad \textit{for all } k = 1, \ldots, m. \tag{2}$$

Proof. The proof is constructive, that is, we will actually construct the vectors $e_1, \ldots, e_m$ with the desired properties. Since $(v_1, \ldots, v_m)$ is linearly independent, $v_k \neq 0$ for all $k = 1, 2, \ldots, m$. Set $e_1 = \frac{v_1}{\|v_1\|}$. This is a vector of norm 1 and satisfies (2) for $k = 1$. Next set

$$e_2 = \frac{v_2 - \langle v_2, e_1 \rangle e_1}{\|v_2 - \langle v_2, e_1 \rangle e_1\|}.$$

This is in fact the normalized version of the orthogonal decomposition (1)

$$w = v_2 - \langle v_2, e_1 \rangle e_1$$

where $w \perp v_2$. Note that $\|e_2\| = 1$ and $\mathrm{span}(e_1, e_2) = \mathrm{span}(v_1, v_2)$.

Now suppose $e_1, \ldots, e_{k-1}$ have been constructed such that $(e_1, \ldots, e_{k-1})$ is an orthonormal list and $\mathrm{span}(v_1, \ldots, v_{k-1}) = \mathrm{span}(e_1, \ldots, e_{k-1})$. Then define

$$e_k = \frac{v_k - \langle v_k, e_1 \rangle e_1 - \langle v_k, e_2 \rangle e_2 - \cdots - \langle v_k, e_{k-1} \rangle e_{k-1}}{\|v_k - \langle v_k, e_1 \rangle e_1 - \langle v_k, e_2 \rangle e_2 - \cdots - \langle v_k, e_{k-1} \rangle e_{k-1}\|}.$$

Since $(v_1, \ldots, v_k)$ is linearly independent we know that $v_k \notin \mathrm{span}(v_1, \ldots, v_{k-1})$. Hence also $v_k \notin \mathrm{span}(e_1, \ldots, e_{k-1})$. Hence the norm in the definition of e_k is not zero and hence e_k is well-defined (we are not dividing by zero). Also, a vector divided by its norm has norm 1.

Hence $\|e_k\| = 1$. Furthermore

$$\begin{aligned}\langle e_k, e_i \rangle &= \left\langle \frac{v_k - \langle v_k, e_1\rangle e_1 - \langle v_k, e_2\rangle e_2 - \cdots - \langle v_k, e_{k-1}\rangle e_{k-1}}{\|v_k - \langle v_k, e_1\rangle e_1 - \langle v_k, e_2\rangle e_2 - \cdots - \langle v_k, e_{k-1}\rangle e_{k-1}\|}, e_i \right\rangle \\ &= \frac{\langle v_k, e_i\rangle - \langle v_k, e_i\rangle}{\|v_k - \langle v_k, e_1\rangle e_1 - \langle v_k, e_2\rangle e_2 - \cdots - \langle v_k, e_{k-1}\rangle e_{k-1}\|} = 0\end{aligned}$$

for $1 \leq i < k$. Hence $(e_1, \ldots, e_k)$ is orthonormal.

From the definition of e_k we see that $v_k \in \text{span}(e_1, \ldots, e_k)$ so that $\text{span}(v_1, \ldots, v_k) \subset \text{span}(e_1, \ldots, e_k)$. Since both lists $(e_1, \ldots, e_k)$ and $(v_1, \ldots, v_k)$ are linearly independent, they must span subspaces of the same dimension and therefore are the same subspace. Hence (2) holds. □

Example 4. Take $v_1 = (1, 1, 0)$ and $v_2 = (2, 1, 1)$ in $\mathbb{R}^3$. The list (v_1, v_2) is linearly independent (check!). Then

$$e_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}}(1, 1, 0).$$

Next

$$e_2 = \frac{v_2 - \langle v_2, e_1\rangle e_1}{\|v_2 - \langle v_2, e_1\rangle e_1\|}.$$

The inner product is $\langle v_2, e_1\rangle = \frac{1}{\sqrt{2}}\langle (1, 1, 0), (2, 1, 1)\rangle = \frac{3}{\sqrt{2}}$, so that

$$u_2 = v_2 - \langle v_2, e_1\rangle e_1 = (2, 1, 1) - \frac{3}{2}(1, 1, 0) = \frac{1}{2}(1, -1, 2).$$

Calculating the norm of u_2 we obtain $\|u_2\| = \sqrt{\frac{1}{4}(1 + 1 + 4)} = \frac{\sqrt{6}}{2}$. Hence normalizing this vector we obtain

$$e_2 = \frac{u_2}{\|u_2\|} = \frac{1}{\sqrt{6}}(1, -1, 2).$$

The list (e_1, e_2) is therefore orthonormal with the same span as (v_1, v_2).

Corollary 11. *Every finite-dimensional inner product space has an orthonormal basis.*

Proof. Let $(v_1, \ldots, v_n)$ be a basis for V. This list is linearly independent and spans V. Apply the Gram-Schmidt procedure to this list to obtain an orthonormal list $(e_1, \ldots, e_n)$ which still spans V. By Proposition 8 this list is linearly independent and hence a basis of V. □

Corollary 12. *Every orthonormal list of vectors in V can be extended to an orthonormal basis of V.*

Proof. Let $(e_1, \ldots, e_m)$ be an orthonormal list of vectors in V. By Proposition 8 this list is linearly independent and hence can be extended to a basis $(e_1, \ldots, e_m, v_1, \ldots, v_k)$ of V

by the Basis Extension Theorem. Now apply the Gram-Schmidt procedure to obtain a new orthonormal basis $(e_1, \ldots, e_m, f_1, \ldots, f_k)$. The first m vectors do not change since they already are orthonormal. The list still spans V and is linearly independent by Proposition 8 and therefore forms a basis. □

Recall that we proved that for a complex vector space V there is always a basis with respect to which a given operator $T \in \mathcal{L}(V, V)$ is upper-triangular. We would like to extend this result to require the additional property of orthonormality.

Corollary 13. *Let V be an inner product space over $\mathbb{F}$ and $T \in \mathcal{L}(V, V)$. If T is upper-triangular with respect to some basis, then T is upper-triangular with respect to some orthonormal basis.*

Proof. Let $(v_1, \ldots, v_n)$ be a basis of V with respect to which T is upper-triangular. Apply the Gram-Schmidt procedure to obtain an orthonormal basis $(e_1, \ldots, e_n)$. Note that

$$\operatorname{span}(e_1, \ldots, e_k) = \operatorname{span}(v_1, \ldots, v_k) \quad \text{for all } 1 \leq k \leq n.$$

We proved before that T is upper-triangular with respect to a basis $(v_1, \ldots, v_n)$ if and only if $\operatorname{span}(v_1, \ldots, v_k)$ is invariant under T for all $1 \leq k \leq n$. Since the span has not changed by going to the basis $(e_1, \ldots, e_n)$, T is still upper-triangular even for this new basis. □

6 Orthogonal projections and minimization problems

Definition 7. Let V be a finite-dimensional inner product space and $U \subset V$ a subset of V. Then the **orthogonal complement** of U is the set

$$U^\perp = \{v \in V \mid \langle u, v \rangle = 0 \text{ for all } u \in U\}.$$

Note that in fact $U^\perp$ is always a subspace of V (as you should check!) and

$$\{0\}^\perp = V, \qquad V^\perp = \{0\}.$$

If $U_1 \subset U_2$, then $U_2^\perp \subset U_1^\perp$.

Furthermore, if $U \subset V$ is not only a subset, but a subspace then we will now show that

$$\begin{aligned} V &= U \oplus U^\perp \\ U^{\perp\perp} &= U. \end{aligned}$$

Theorem 14. *If $U \subset V$ is a subspace of V, then $V = U \oplus U^\perp$.*

Proof. We need to show that

1. $V = U + U^\perp$.
2. $U \cap U^\perp = \{0\}$.

To show 1, let $(e_1, \ldots, e_m)$ be an orthonormal basis of U. Then for all $v \in V$ we can write

$$v = \underbrace{\langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m}_{u} + \underbrace{v - \langle v, e_1\rangle e_1 - \cdots - \langle v, e_m\rangle e_m}_{w}. \tag{3}$$

The vector $u \in U$ and

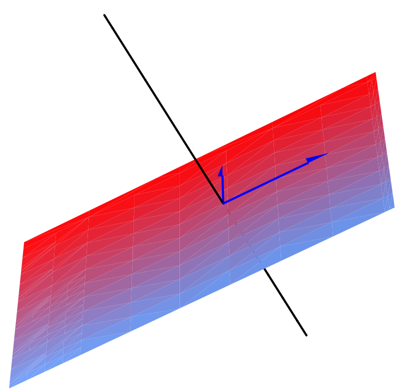
$$\langle w, e_j\rangle = \langle v, e_j\rangle - \langle v, e_j\rangle = 0 \qquad \text{for all } j = 1, 2, \ldots, m$$

since $(e_1, \ldots, e_m)$ is an orthonormal list of vectors. Hence $w \in U^\perp$. This implies that $V = U + U^\perp$.

To prove 2, let $v \in U \cap U^\perp$. Then v has to be orthogonal to every vector in U, in particular to itself, so that $\langle v, v\rangle = 0$. However, this implies $v = 0$, so that $U \cap U^\perp = \{0\}$. □

Example 5. $\mathbb{R}^2$ is the direct sum of two orthogonal lines and $\mathbb{R}^3$ is the direct sum of a plane and a line orthogonal to this plane. For example

$$\begin{aligned}
\mathbb{R}^2 &= \{(x, 0) \mid x \in \mathbb{R}\} \oplus \{(0, y) \mid y \in \mathbb{R}\}\\
\mathbb{R}^3 &= \{(x, y, 0) \mid x, y \in \mathbb{R}\} \oplus \{(0, 0, z) \mid z \in \mathbb{R}\}.
\end{aligned}$$



Theorem 15. *If $U \subset V$ is a subspace of V, then $U = (U^\perp)^\perp$.*

Proof. First we show that $U \subset (U^\perp)^\perp$. Let $u \in U$. Then for all $v \in U^\perp$ we have $\langle u, v\rangle = 0$. Hence $u \in (U^\perp)^\perp$ by the definition of $(U^\perp)^\perp$.

Next we show that $(U^\perp)^\perp \subset U$. Suppose $0 \neq v \in (U^\perp)^\perp$ such that $v \notin U$. Decompose v according to Theorem 14

$$v = u_1 + u_2 \in U \oplus U^\perp,$$

where $u_1 \in U$ and $u_2 \in U^\perp$. Then $u_2 \neq 0$ since $v \notin U$. Furthermore, $\langle u_2, v\rangle = \langle u_2, u_2\rangle \neq 0$. But then v is not in $(U^\perp)^\perp$ which contradicts our initial assumption. Hence we must have $(U^\perp)^\perp \subset U$. □

By Theorem 14 we have the decomposition $V = U \oplus U^\perp$ for every subspace $U \subset V$. Hence we can define the **orthogonal projection** P_U of V onto U as follows. Every $v \in V$ can be uniquely written as $v = u + w$ where $u \in U$ and $w \in U^\perp$. Define

$$\begin{aligned} P_U : V &\to V \\ v &\mapsto u. \end{aligned}$$

Clearly, P_U is a projection operator since $P_U^2 = P_U$. It also satisfies

$$\begin{aligned} \operatorname{range} P_U &= U \\ \operatorname{null} P_U &= U^\perp \end{aligned}$$

so that $\operatorname{range} P_U \perp \operatorname{null} P_U$. Therefore P_U is called an orthogonal projection.

The decomposition of a vector $v \in V$ into a piece in U and a piece in $U^\perp$ as given in (3) yields the following formula for P_U

$$P_U v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m,$$

where $(e_1, \ldots, e_m)$ is an orthonormal basis of U.

Let us now apply the inner product to the following minimization problem. Given a subspace $U \subset V$ and a vector $v \in V$, find the vector $u \in U$ that is closest to the vector v, that is, such that $\|v - u\|$ is smallest. The next proposition shows that $P_U v$ is the closest point in U to the vector v and that this minimum is in fact unique.

Proposition 16. *Let $U \subset V$ be a subspace of V and $v \in V$. Then*

$$\|v - P_U v\| \le \|v - u\| \qquad \textit{for every } u \in U.$$

Furthermore, equality holds if and only if $u = P_U v$.

Proof. Let $u \in U$ and set $P := P_U$ for short. Then

$$\begin{aligned} \|v - Pv\|^2 &\le \|v - Pv\|^2 + \|Pv - u\|^2 \\ &= \|(v - Pv) + (Pv - u)\|^2 = \|v - u\|^2, \end{aligned}$$

where the second line follows from Pythagoras' Theorem 3 since $v - Pv \in U^{\perp}$ and $Pv - u \in U$. Equality only holds if $\|Pv - u\|^2 = 0$ which is equivalent to $Pv = u$. $\square$