We will now come back to our original aim: Given a vector space \( V \), a subspace \( W \), and \( v \in V \). Find the vector \( w \in W \) which is closest to \( v \).

First let us clarify what "closest to" means. The tool to measure distance in the norm, so we want \( \| v - w \| \) to be as small as possible. Thus our problem is: Find a vector \( w \in V \) such that
\[
\| u - w \| \leq \| u - v \|
\]
for all \( v \in W \).

Now, let us recall that if \( W = \mathbb{R} w \) is a line, then the vector \( w \) on the line \( W \) is the one with the property that \( v - w \perp W \). We will start by showing that this will always be the case.
Theorem Let \( V \) be a vector space with inner product \((\cdot, \cdot)\). Let \( W \subseteq V \) be a subspace and \( v \in V \). If \( v - w \perp W \) then \( \| v - w \| \leq \| v - u \| \) for all \( u \in W \) and \( \| v - w \| = \| v - u \| \) if and only if \( w = u \). Thus \( w \) is the member of \( W \) closest to \( v \).

Proof. First we remark that \( \| v - w \| \leq \| v - u \| \) if and only if \( \| v - w \|^2 \leq \| v - u \|^2 \). Now we simply calculate

\[
\| v - u \|^2 = \| (v - w) + (w - u) \|^2
\]

\[
= \| v - w \|^2 + \| w - u \|^2 \quad \text{because } v - w \perp W \text{ and } w - u \in W
\]

(\(*\) \( \geq \| v - w \|^2 \)) \( \leq \) because \( \| w - u \|^2 \geq 0 \).

So \( \| v - u \| \geq \| v - w \| \). If \( \| v - u \|^2 = \| v - w \|^2 \), then we see – using (\(*\)) – that \( \| w - u \|^2 = 0 \), or \( w = u \). So \( \| v - w \| = \| v - u \| \) if \( u = w \), we have shown that the statement is correct. \( \square \)
Theorem: Let $V$ be a vector space with inner product $(\cdot, \cdot)$. Let $W \subseteq V$ be a subspace, and $v \in V$. If $w \in W$ is closest to $v$ then $v - w \perp W$.

Proof: We know that $\|v - w\|^2 \leq \|v - u\|^2$ for all $u \in W$. Therefore the function $f: \mathbb{R} \to \mathbb{R}$

$$F(t) = \|v - w + tx\|^2 \quad (x \in W)$$

has a minimum at $t = 0$. We have

$$F(t) = (v - w + tx, v - w + tx)$$

$$= (v - w, v - w) + t(v - w, x)$$

$$+ t(x, v - w) + t^2 (x, x)$$

$$= \|v - w\|^2 + 2t(v - w, x) + t^2 \|x\|^2 .$$

Therefore

$$0 = F'(0) = 2(v - w, x) .$$

As $x \in W$ was arbitrary, it follows that $v - w \perp W$.

Our task is now to construct the vector $w$ such that $v - w \perp W$. The idea, how to do that comes from the Gram-Schmidt.
Let \( W = \mathbf{1} \mathbf{u} \) and \( \mathbf{v} \in V \). The Gram-Schmidt applied to \( \mathbf{u} \) and \( \mathbf{v} \) shows, that

\[
\mathbf{v} = \frac{(\mathbf{v}, \mathbf{u})}{\|\mathbf{u}\|^2} \mathbf{u} \perp W
\]

so \( \mathbf{w} = \frac{(\mathbf{v}, \mathbf{u})}{\|\mathbf{u}\|^2} \mathbf{u} \) is the vector (point) on the line \( W \) closest to \( \mathbf{v} \). What if the dimension of \( W \) is greater than one? Assume that we have found an orthogonal basis \( \mathbf{v}_1, \ldots, \mathbf{v}_n \) for \( W \). Applying the Gram-Schmidt to the vectors \( \mathbf{v}_1, \ldots, \mathbf{v}_n \) shows that

\[
\mathbf{w} - \sum_{j=1}^{n} \frac{(\mathbf{w}, \mathbf{v}_j)}{\|\mathbf{v}_j\|^2} \mathbf{v}_j
\]

is orthogonal to each one of the vectors \( \mathbf{v}_1, \ldots, \mathbf{v}_n \), and hence because the inner product is linear in each factor at the time-orthogonal to all vectors of the form \( c_1 \mathbf{v}_1 + \cdots + c_n \mathbf{v}_n = \sum_{j=1}^{n} c_j \mathbf{v}_j \). But each vector of \( W \) can be written in this way. Thus our vector \( \mathbf{w} \) closest to \( \mathbf{v} \) is given by

\[
\mathbf{w} = \sum_{j=1}^{n} \frac{(\mathbf{v}, \mathbf{v}_j)}{\|\mathbf{v}_j\|^2} \mathbf{v}_j
\]
Let us look at another motivation: We know that \( w \in W \). Hence, if \( v_1, \ldots, v_n \in W \) is a basis, we can find scalars \( c_1, \ldots, c_n \in \mathbb{R} \) such that

\[
  w = \sum_{k=1}^{m} c_k v_k
\]

So what are those numbers? As \( v - w \perp v_j \) for \( j = 1, \ldots, m \) and \( v_{k \neq j} \perp v_j \) for \( k \neq j \) we get:

\[
  0 = (v_j, v_j) - (w, v_j) = (v_j, v_j) - \sum_{k \neq j} c_k (v_k, v_j)
\]

So solving for \( c_j \) we get

\[
  c_j = \frac{(v_j, v_j)}{\| v_j \|^2}
\]

Thus

\[
  w = \sum_{j=1}^{m} \frac{(v_j, v_j)}{\| v_j \|^2} v_j
\]

We collect this in the following theorem:
We will now prove this directly:

**Theorem.** Let $V$ be a vector space with inner product $(\cdot, \cdot)$. Let $W \subset V$ be a subspace and assume that $v_1, \ldots, v_n \in W$ is an orthogonal basis for $W$. For $v \in V$ let $w = \sum_{j=1}^{n} \frac{(v, v_j)}{\|v_j\|^2} v_j \in W$. Then $v - w \perp W$ (or $w$ is the vector in $W$ closest to $v$).

**Proof.** We have

\[
0 = (v - w, v_j) = (v, v_j) - (w, v_j) = (v, v_j) - \sum_{k=1}^{n} \frac{(v, v_k)}{\|v_k\|^2} (v_k, v_j) = (v, v_j) - \frac{(v, v_j)}{\|v_j\|^2} \|v_j\|^2 \left\{ \begin{array}{l} (v, v_j) = 0 \quad \text{if } i = j \\ (v, v_j) = \|v_j\|^2 \quad \text{if } i \neq j \end{array} \right.
\]

\[
= (v, v_j) - (v, v_j) = 0 \Leftrightarrow (v_j, v_j) = \|v_j\|^2.
\]

Hence $v - w \perp v_j$. But, as we just saw, this implies, that $v - w \perp W$ because $v_1, \ldots, v_n$ is a basis.

Let us now look at what we just said from the point of view of linear maps. What is given in the beginning is a vector space with an inner product and a subspace $W$. Then, for each $v \in V$, we associated an unique vector $w \in W$. Thus we got a map.
\[ P: V \rightarrow W, \quad v \mapsto w. \]

We even have an explicit formula for \( P(v) \):

Let \( v_1, \ldots, v_n \) be an orthogonal basis for \( W \), then

\[
P(v) = \sum_{k=1}^{n} \frac{(V, V_k)}{\|V_k\|^2} V_k.
\]

This shows that \( P \) is linear. We saw earlier that if \( v \in W \), then

\[
v = \sum_{k=1}^{n} \frac{(V, V_k)}{\|V_k\|^2} V_k.
\]

So \( P(v) = v \) for all \( v \in V \). In particular we get

**Lemma** \( P^2 = P \).

The map \( P \) is called the \underline{orthogonal projection} onto \( W \).

The projection part comes from \( P^2 = P \) and \underline{orthogonal} from the fact that \( v - P(v) \perp W \).

The result of this discussion is the following:

To find the vector \( w \) closest to \( v \) we have to:

1. Find (if possible) a basis \( u_1, \ldots, u_n \) for \( W \).

2. If this is not an orthogonal basis, then we use
Gram-Schmidt to construct an orthonormal basis \( V_1, \ldots, V_n \).

(3) Then \( W = \sum_{k=1}^{n} \frac{\langle V_i, V_k \rangle}{\|V_k\|^2} V_k \).

**Example**: Let \( W \) be the line \( W = \mathbb{R}(1, 2) \).

Then \( V = (1, 2) \) is a basis (orthogonal!) for \( W \).

It follows, that the orthogonal projection is given by

\[
P(x, y) = \frac{x + 2y}{5} (1, 2).
\]

Let \( (x, y) = (3, 1) \). Then (as an example)

\[
P(3, 1) = (1, 2).
\]

**Example**: Let \( W \) be the line given by \( y = 3x \).

Then \( (1, 3) \in W \) and hence \( W = \mathbb{R}(1, 3) \). It follows that

\[
P(x, y) = \frac{x + 3y}{10} (1, 3).
\]

**Example**: Let \( W \) be the plane generated by the vectors \( (1, 1, 1) \) and \( (1, 0, 1) \). Find the orthogonal projection \( P : \mathbb{R}^3 \to W \).

**Solution**: We notice first that \( \langle (1, 1, 1), (1, 0, 1) \rangle = 2 \neq 0 \) so this is not an orthogonal basis. Using Gram-Schmidt we get:
\[ v_1 = (1, 1, 1) \]

\[ v_2 = (1, 0, 1) - \frac{2}{3} (1, 1, 1) = \left( \frac{1}{3}, -\frac{2}{3}, \frac{1}{3} \right) = \frac{1}{3} (1, -2, 1) \]

To avoid fractions, we can use \((1, -2, 1)\) instead of \(\frac{1}{3} (1, -2, 1)\). Thus, the orthogonal projection is:

\[ P(x, y, z) = \frac{x+y+z}{3} (1, 1, 1) + \frac{x-2y+z}{6} (1, -2, 1) \]

\[ = \left( \frac{2x+2y+2z}{6} + \frac{x-2y+z}{6}, \frac{2x+2y+2z}{6} - 2 \frac{x-2y+z}{6} \right) \]

\[ = \left( \frac{x+z}{2}, \frac{x+z}{2} \right) \]

**Example:** Let \( W \) be the plane \( \{(x, y, z) \in \mathbb{R}^3 : x+y+2z=0\} \).

Find the orthogonal projection \( P : \mathbb{R}^3 \to W \).

**Solution.** We notice that our first step is to find an orthogonal basis for \( W \). The vectors \((1, -1, 0)\) and \((2, 0, -1)\) are in \( W \), but are not orthogonal.

But

\[ (2, 0, -1) - \frac{2}{2} (1, -1, 0) = (1, 1, -1) \in W \]

and is orthogonal to \((1, -1, 0)\). So we get:

\[ P(x, y, z) = \frac{x-y}{2} (1, -1, 0) + \frac{x+y-z}{3} (1, 1, -1) \]

\[ = \left( \frac{5x-y-2z}{6}, \frac{-x+5y-2z}{6}, \frac{-x+y+z}{3} \right) \].