Lecture 8: Orthogonal projections

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.

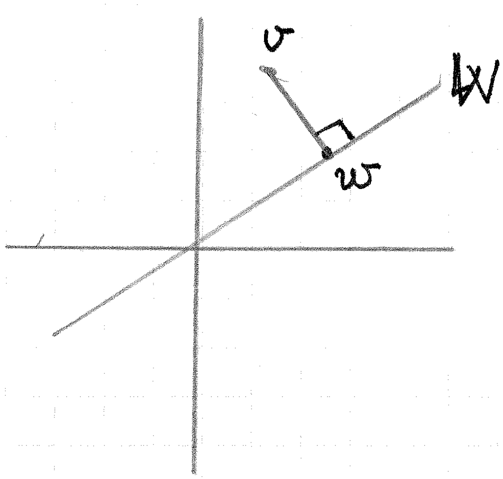
First let us clarify what "closest to" means. The tool to measure distance is the norm, so we want $\|v - w\|$ to be as small as possible. Thus our problem is:

Find a vector $w \in V$ such that

$$\|u - w\| \leq \|u - V\|$$

for all $v \in W$.

Now, let us recall that if $W = \mathbb{R} w_1$ 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 $W \subset V$ 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

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

$$= \|v - w\|^2 + \|w - u\|^2 \leftarrow \text{because } v - w \perp W \text{ and } w - u \in W$$

$$(*) \qquad \geq \|v - w\|^2 \qquad \leftarrow \text{because } \|w - u\|^2 \geq 0.$$

So $\|v - u\| \geq \|v - w\|$. If $\|v - u\|^2 = \|v - w\|^2$, the we see – using $(*)$ – that $\|w - u\|^2 = 0$, or $w = u$. As $\|v - w\| = \|v - u\|$ if $u = w$, we have shown that the statement is correct. ∎

<u>Theorem</u> Let V be a vector space with inner product $(\cdot,\cdot)$. Let $W \subset V$ be a subspace, and $v \in V$. If $w \in W$ is closest to v then $v - w \perp W$.

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

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

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

$$\begin{aligned} F(t) &= (v - w + tx, v - w + tx) \\ &= (v-w, v-w) + t(v-w, x) \\ &\quad + t(x, v-w) + t^2(x,x) \\ &= \|v-w\|^2 + 2t(v-w,x) + t^2\|x\|^2 . \end{aligned}$$

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 $W = \mathbb{R}u$ and $v \in V$. The Gram-Schmidt applied to u and v shows, that

$$v - \frac{(v,u)}{\|u\|^2}\,u \perp W$$

So $w = \dfrac{(v,u)}{\|u\|^2}\,u$ is the vector (point) on the line W closest to v. What if the dimension of W is greater than one? Assume that we have found an orthogonal basis $v_1, \ldots, v_n$ for W. Applying the Gram-Schmidt to the vectors $v_1, \ldots, v_n, v$ shows that

$$v - \sum_{j=1}^{m} \frac{(v, v_j)}{\|v_j\|^2}\,v_j$$

is orthogonal to each one of the vectors $v_1, \ldots, v_n$, and hence - because the inner product is linear in each factor at the time - orthogonal to all vectors of the form $c_1 v_1 + \cdots + c_n v_n = \sum_{j=1}^{m} c_j v_j$. But each vector of W can be written in this way. Thus our vector w closest to v is given by

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

Let us look at another motivation; We know that $w \in W$. Hence, if $v_1, \dots, v_n \in W$ is a basis, we can find scalars $c_1, \dots, c_n \in \mathbb{R}$ such that

$$w = \sum_{k=1}^{m} c_k v_k .$$

So what are those numbers? As $v - w \perp v_j$ for $j = 1, \dots, m$ and $v_k \perp v_j$ for $k \neq j$ we get:

$$\begin{aligned} 0 &= (v, v_j) - (w, v_j) \\ &= (v, v_j) - \sum_{k=1}^{m} c_k (v_k, v_j) \\ &= (v, v_j) - c_j \|v_j\|^2 . \end{aligned}$$

Solving for c_j we get

$$c_j = \frac{(v, v_j)}{\|v_j\|^2} .$$

Thus

$$w = \sum_{j=1}^{m} \frac{(v, v_j)}{\|v_j\|^2} v_j .$$

We collect this in the following theorem:

We will now prove this directly:

<u>Theorem</u> Let V be a vector space with inner product $(\cdot\,,\cdot)$. Let $W \subset V$ be a subspace and assume that $v_1, \dots, v_n \in W$ 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).

<u>Proof.</u> We have

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

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

Let us now look at what we just did 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

$$P : V \to W, \; v \mapsto w.$$

We even have an explicite formula for $P(v)$:

Let (if possible) $v_1, \dots, v_n$ be an orthogonal basis for W, then

$$P(v) = \sum_{k=1}^{m} \frac{(v, v_k)}{\|v_k\|^2} v_k .$$

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

$$v = \sum_{k=1}^{m} \frac{(v, v_k)}{\|v_k\|^2} v_k$$

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

Lemma $P^2 = P$.

The map P is called the orthogonal projection onto W.

The projection part comes from $P^2 = P$ and 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 $u_1, \dots, u_l$ 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 $w = \sum_{k=1}^{n} \frac{(v, v_k)}{\|v_k\|^2} v_k$.

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 $((1,1,1),(1,0,1)) = 2 \neq 0$ so this is not an orthogonal basis. Using Gram-Schmidt we get:

$V_1 = (1,1,1)$

$V_2 = (1,0,1) - \frac{2}{3}(1,1,1) = (\frac{1}{3}, -\frac{2}{3}, \frac{1}{3}) = \frac{1}{3}(1,-2,1).$

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

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

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

$$= \left(\frac{x+z}{2}, y, \frac{x+z}{2}\right).$$

<u>Example</u>: Let W be the the plane $\{(x,y,z) \in \mathbb{R}^3 : x+y+2z=0\}$. Find the orthogonal projection $P: \mathbb{R}^3 \to W$.

<u>Solution</u> 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:

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

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