Lecture 8

Orthogonal Projections

8.1 Introduction

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

First let us clarify what the "closest" 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 $w \in W$ such that

$$\|v - w\| \leq \|v - u\|$$

for all $u \in W$.

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 is always the case.

8.2 $w \in W$ is closest to v *iff* $v - w \perp W$

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

$$\begin{aligned} \|v-u\|^2 &= \|(v-w)+(w-u)\|^2 \\ &= \|v-w\|^2 + \|w-u\|^2 \\ &\quad \text{because } v-w \perp W \text{ and } w-u \in W \\ (*) \quad &\geq \|v-w\|^2 \quad \text{because } \|w-u\|^2 \geq 0 \end{aligned}$$

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

Theorem. 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 the 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 $f : \mathbb{R} \longrightarrow \mathbb{R}$

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

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 $v - w \perp W$. □

8.2.1 Construction of w

Our task now is to construct the vector w such that $v - w \perp W$. The idea is to use Gram-Schmidt orthogonalization.
Let $W = \mathbb{R}u$ and $v \in V$. Applying Gram-Schmidt to u and v gives:

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

So that $w = \frac{(v,u)}{\|u\|^2} u$ is the vector (point) on the line W closest to v.

What if the dimension of W is greater than one? Let $v_1, \ldots, v_n$ be an orthogonal basis for W. Applying the Gram-Schmidt to the vectors $v_1, \ldots, v_n, v$ shows that

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

is orthogonal to each one of the vectors $v_1, \ldots, v_n$. Since

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

is orthogonal to v_j for all j, it is orthogonal to any linear combination of them $c_1 v_1 + \ldots + c_n v_n = \sum_{j=1}^{n} c_j v_j$, and hence it is orthogonal to W. Therefore our vector w closest to v is given by

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

Let us look at another motivation; Let $w \in W$ be the closest to v and let $v_1, \ldots, v_n$ be a basis for W. Then there are scalars $c_1, \ldots, c_n \in \mathbb{R}$ such that

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

So what are these scalars? As $v - w \perp v_j$ for $j = 1, \ldots, n$ and $v_k \perp v_j$ for

$k \neq j$ we get:

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

Solving for c_j we get

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

Thus

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

8.3 The main theorem

We collect the results of the above computations in the following (main)theorem:

Theorem. Let V be a vector space with inner product $(\cdot, \cdot)$. Let $W \subset V$ be a subspace and assume that $\{v_1, \ldots, v_n\}$ 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 equivalently, w is the vector in W closest to v).

Proof. We have

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

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

8.4 Orthogonal projections

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 a unique vector $w \in W$. Thus we got a map

$$P : V \longrightarrow W, \quad v \mapsto w$$

We even have an explicit formula for $P(v)$: Let (if possible) $v_1, \ldots, v_n$ 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 showed 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 $v \in W$. 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 $v - P(v) \perp W$.

8.5 Summary

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 use Gram-Schmidt to construct an orthogonal basis $v_1, \ldots, v_n$.

3. Then $w = \sum_{k=1}^{n} \frac{(v,v_k)}{\|v_k\|^2} v_k$.

8.6 Examples

Example. Let W be the line $W = \mathbb{R}(1,2)$. Then $u = (1,2)$ 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

$$P(3,1) = (1,2).$$

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 $P : \mathbb{R}^3 \longrightarrow W$.

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 $(1,-2,1)$ instead of $\frac{1}{3}(1,-2,1)$. Thus the orthogonal projection is:

$$\begin{aligned} P(x,y,z) &= \frac{x+y+z}{3}(1,1,1) + \frac{x-2y+z}{6}(1,-2,1) \\ &= \Bigg(\frac{2x+2y+2z}{6} + \frac{x-2y+z}{6}, \\ &\quad \frac{2x+2y+2z}{6} - 2\frac{x-2y+z}{6}, \\ &\quad \frac{2x+2y+2z}{6} + \frac{x-2y+z}{6} \Bigg) \\ &= \left(\frac{x+z}{2}, y, \frac{x+z}{2} \right). \end{aligned}$$

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

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.We have

$$(2,0,-1) - \tfrac{2}{2}(1,-1,0) = (1,1,-1) \in W$$

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

$$\begin{aligned} P(x,y,z) &= \frac{x-y}{2}(1,-1,0) + \frac{x+y-z}{3}(1,1,-1) \\ &= \left(\frac{5x-y-2z}{6}, \frac{-x+5y-2z}{6}, \frac{-x-y+z}{3}\right). \end{aligned}$$

8.7 Exercises

1. Let $V \subset \mathbb{R}^2$ be the line $V = \mathbb{R}(1,-1)$.
   (a) Write a formula for the orthogonal projection $P : \mathbb{R}^2 \to V$.
   (b) What is: i) $P(1,1)$, ii) $P(2,1)$, iii) $P(2,-2)$?
2. Let $W \subset \mathbb{R}^3$ be the plane
   $$W = \{(x,y,z) \in \mathbb{R}^3 : \; x - 2y + z = 0\}.$$
   (a) Find the orthogonal projection $P : \mathbb{R}^3 \to W$.
   (b) What is: i) $P(1,1,2)$, ii) $P(1,-2,1)$, iii) $P(2,1,1)$?
3. Let $W \subset \mathbb{R}^3$ be the plane generated by the vectors $(1,1,1)$ and $(1,-1,1)$.
   (a) Find the orthogonal projection $P : \mathbb{R}^3 \to W$.
   (b) What is: i) $P(1,1,2)$, ii) $P(2,0,1)$?
4. Let W be the space of continuous functions on $[0,1]$ generated by the constant function 1 and x. Thus $W = \{a_0 + a_1 x : \; a_0, a_1 \in \mathbb{R}\}$. Find the orthogonal projection of the following functions onto W:
   i) $P(x^2)$, ii) $P(e^x)$, iii) $P(1+x^2)$.
5. Let W be the space of piecewise continuous functions on $[0,1]$ generated by $\chi_{[0,1/2)}$ and $\chi_{[1/2,1)}$. Find orthogonal projections of the following functions onto W:
   i) $P(x)$, ii) $P(x^2)$, iii) $P(\chi_{[0,3/4]})$.