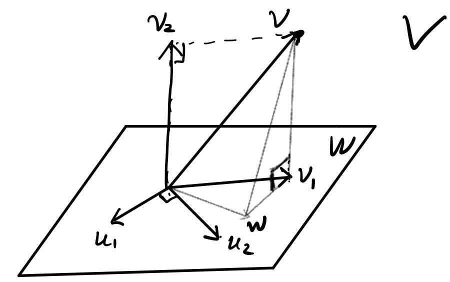
$V, \langle \cdot, \cdot \rangle$: Inner product space

$W \subseteq V$ a subspace.



Find the closest vector $v_1 \in W$ to v.

Thm: 1. There exists a unique vector $v_1 \in W$ that is closest to v.

2. The closest vector $v_1 \in W$ to v satisfies: $v - v_1 \perp W$

This means: $\langle v - v_1, w \rangle = 0$ for any $w \in W$.

Proof: Choose any basis $\{w_1, w_2, \cdots, w_r\}$ for W. We can use the Gram-Schmidt orthogonalization process to get an o.n.b. $\{u_1, u_2, \cdots, u_r\}$.

Set $$\underset{\text{Proj}_W v}{\overset{\|}{v_1}} = \langle v, u_1 \rangle u_1 + \langle v, u_2 \rangle u_2 + \cdots + \langle v, u_r \rangle u_r = \sum_{i=1}^{r} \langle v, u_i \rangle u_i$$

We will prove v_1 is the closest vector to v in W.

- First we show $v_2 = v - v_1 \perp W$. For any u_j, $1 \le j \le r$,

$$\langle v_2, u_j \rangle = \langle v, u_j \rangle - \sum_{i=1}^{r} \langle v, u_i \rangle \langle u_i, u_j \rangle = \langle v, u_j \rangle - \langle v, u_j \rangle = 0$$

$\Rightarrow \forall\ w = c_1 u_1 + \cdots + c_r u_r \in W$, $\langle v_2, w \rangle = 0$.

- Next we show that v_1 is the closest: For any $w \in W$, $v_1 - w \in W$

$\Rightarrow v_2 \perp v_1 - w \Rightarrow \langle v - v_1, v_1 - w \rangle = 0$

Pythagorean thm

$\Longrightarrow \underset{\underset{(v-v_1)+(v_1-w)}{\shortparallel}}{\|v-w\|^2} = \|v-v_1\|^2 + \|v_1-w\|^2 \geq \|v-v_1\|^2$

$\Rightarrow \|v-w\| \geq \|v-v_1\|$. and equality holds only when $v_1 = w$.

$\Downarrow$

- Finally we show that v_1 is unique:

  If $v_1' \in W$ is also closest, then $\|v-v_1'\| = \|v-v_1\| \Rightarrow v_1' = v_1$. ∎

The proof shows that to find $\mathrm{Proj}_W v$. We can

- Find an o.n.b. $\{u_1, u_2, \cdots, u_r\}$ for W
- Calculate $\mathrm{Proj}_W v = \sum_{i=1}^{r} \langle v, u_i \rangle u_i$.

<u>Ex</u>: $V = P_2(\mathbb{R})$. $\langle f, g \rangle = \int_0^2 f(t) g(t) dt$

$W = P_1(\mathbb{R})$. $h(t) = 2 - x + x^2$.

$W = P_1(\mathbb{R}) = \mathrm{Span}\{\underset{w_1}{\underset{\shortparallel}{1}}, \underset{w_2}{\underset{\shortparallel}{x}}\}$

$v_1 = w_1 = 1$. $\|v_1\|^2 = \int_0^2 1 \cdot 1 \, dt = 2$

$v_2 = w_2 - \frac{\langle w_2, v_1 \rangle}{\|v_1\|^2} v_1$ $\quad \langle w_2, v_1 \rangle = \int_0^2 x \cdot 1 \, dx = \frac{x^2}{2}\Big|_0^2 = 2$.

$v_2 = x - \frac{2}{2} \cdot 1 = x - 1$

$\|v_2\|^2 = \int_0^2 (x-1)^2 dx = \frac{(x-1)^3}{3}\Big|_0^2 = \frac{2}{3}$

$$\Rightarrow u_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}}, \quad u_2 = \frac{v_2}{\|v_2\|} = \frac{x-1}{\sqrt{\frac{2}{3}}} = \sqrt{\frac{3}{2}}(x-1).$$

$$\langle h, u_1\rangle = \int_0^2 (2-t+t^2)\cdot\frac{1}{\sqrt{2}}\,dt = \frac{1}{\sqrt{2}}\cdot\left(2t-\frac{t^2}{2}+\frac{t^3}{3}\right)\Big|_0^2 = \frac{1}{\sqrt{2}}\left(4-2+\frac{8}{3}\right) = \frac{1}{\sqrt{2}}\frac{14}{3}$$

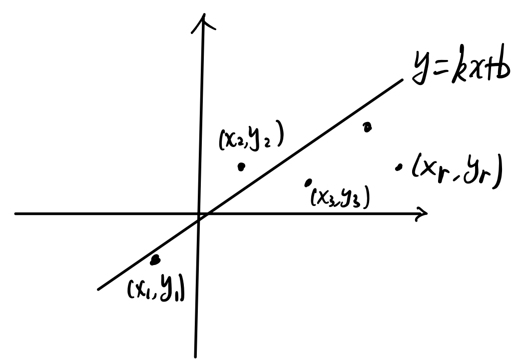
$$\langle h, u_2\rangle = \int_0^2 (2-t+t^2)\cdot\sqrt{\frac{3}{2}}(t-1)\,dt \underset{s=(t-1)}{=} \sqrt{\frac{3}{2}}\int_{-1}^1 \underbrace{(2-(s+1)+(s+1)^2)\cdot s}_{\overset{\parallel}{(1-s+s^2+2s+1)\cdot s}}\,ds$$

$$= \sqrt{\frac{3}{2}}\cdot\int_{-1}^1 (s+s^2+s^3)\,ds = \sqrt{\frac{3}{2}}\cdot 2\cdot\frac{s^3}{3}\Big|_0^1 = \sqrt{\frac{3}{2}}\cdot\frac{2}{3}.$$

$$\Rightarrow \mathrm{Proj}_W h = \langle h, u_1\rangle u_1 + \langle h, u_2\rangle u_2$$

$$= \frac{1}{\sqrt{2}}\frac{14}{3}\frac{1}{\sqrt{2}} + \sqrt{\frac{3}{2}}\cdot\frac{2}{3}\cdot\sqrt{\frac{3}{2}}(x-1) = \frac{7}{3} + x - 1 = x + \frac{4}{3}.$$

Application: Optimal linear approximation



Find (k, b) s.t. the Error E is smallest:

$$E = (kx_1+b-y_1)^2 + \cdots + (kx_r+b-y_r)^2$$

$$= \Bigg\| \underbrace{\begin{pmatrix} kx_1+b \\ \vdots \\ kx_r+b \end{pmatrix} - \begin{pmatrix} y_1 \\ \vdots \\ y_r \end{pmatrix}}_{\overset{\parallel}{\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & 1 \end{pmatrix}\begin{pmatrix} b \\ k \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_r \end{pmatrix}}} \Bigg\|^2$$

Let $A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & 1 \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_r \end{pmatrix}$

The problem becomes:

Find $v_1 = \begin{pmatrix} b \\ k \end{pmatrix} \in \mathbb{R}^2$ s.t. $\|y - Av_1\|$ is smallest.

In particular $Av_1 = \mathrm{Proj}_W y$ where $W = R(A)$ is the range of A
$\parallel$
Column space of A.

We can apply the previous method to find $y_1 = \mathrm{Proj}_W y$ and then solve for v_1. But there is another more explicit formula.

To derive this, recall that $y_1 = Av_1 = \mathrm{Proj}_W y$ satisfies $y_1 \perp W$:

$\Rightarrow \quad \forall\ Av \in R(A)$ with $v \in \mathbb{R}^2$, $\langle Av, y - y_1 \rangle = 0$.

Note for column vectors $y, y' \in \mathbb{R}^n$, $\langle y, y' \rangle = y^T y' \in \mathbb{R}$.

So $\quad 0 = \langle Av, y - Av_1 \rangle = (Av)^T (y - Av_1) = v^T A^T (y - Av_1)$

$$= v^T (A^T y - A^T A v_1) = \langle v, A^T y - A^T A v_1 \rangle = 0 \quad \forall v \in \mathbb{R}^2$$

$\Rightarrow \quad A^T y - A^T A v_1 = 0 \Rightarrow v_1 = (A^T A)^{-1} A^T y$.

In general, let $A : \mathbb{R}^n \to \mathbb{R}^m$ be an $m \times n$ matrix.

Assume $\mathrm{rank}(A) = n$. $\Rightarrow N(A) = \{0\} \subset \mathbb{R}^n$.

Then for any $y \in \mathbb{R}^m$, $\mathrm{Proj}_{R(A)} y = A v_1$ with

$$v_1 = (A^T A)^{-1} A^T y \Rightarrow A v_1 = A (A^T A)^{-1} A^T y.$$

(Lemma: $N(A^T A) = N(A)$.).