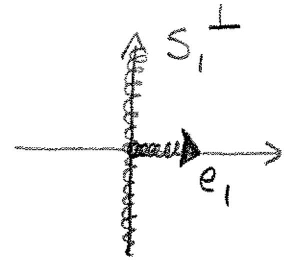
# Lecture 34: Orthogonal complements and projections (§6.2 and §6.3)

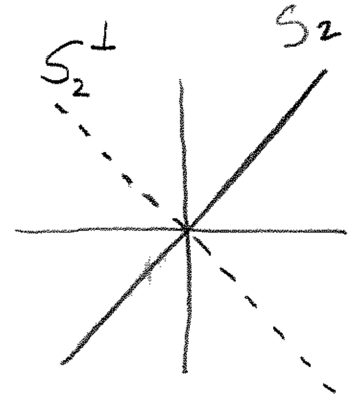
**Def:** Suppose $S$ is a nonempty subset of an inner product space $V$. The <u>orthogonal complement</u> of $S$ is $S^\perp = \{x \in V \mid \langle x, y\rangle = 0 \text{ for all } y \in S\}$.

**Ex:** ($\mathbb{R}^2$, dot prod)




$S_1 = \{e_1\} \Rightarrow S_1^\perp = \text{span}(\{e_2\})$

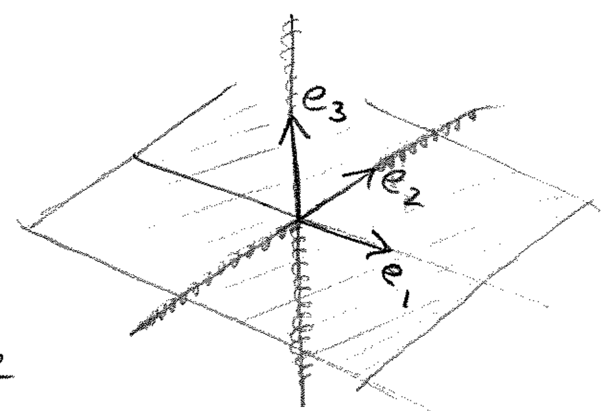
$S_2 = \text{span}(\{(1,1)\}) \Rightarrow S_2^\perp = \text{span}(\{(-1,1)\})$




**Ex:** ($\mathbb{R}^3$, dot prod)

$S_1 = \{z\text{-axis}\} \qquad S_1^\perp = \{xy\text{-plane}\}$

$S_2 = \{e_1, e_3\} \qquad S_2^\perp = \{y\text{-axis}\}$

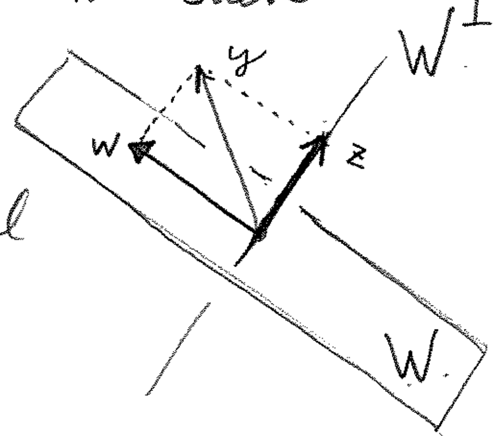



**Note:** $S^\perp$ is always a subspace since if $x_1, x_2 \in S^\perp$ then $\langle c x_1 + x_2, y\rangle$

$= c\langle x_1, y\rangle + \langle x_2, y\rangle = c \cdot 0 + 0 = 0$ for all $y \in S$.

Also $S \cap S^\perp$ contains at most the zero vector.

Thm: Suppose $W$ is a finite-dim'l subspace of an inner product space $V$. For each $y \in V$ there are unique vectors $w \in W$ and $z \in W^{\perp}$ such that $y = w + z$. Moreover, if $\{u_1, u_2, \dots, u_k\}$ is an orthonormal basis for $W$ then




$$w = \sum_{i=1}^{k} \langle y, u_i \rangle u_i$$

Def: This $w$ is called the orthogonal projection of $y$ onto $W$, and gives a linear transformation

proj: $V \longrightarrow W$.

Proof of Thm: Set $w = \sum_{i=1}^{k} \langle y, u_i \rangle u_i$ and $z = y - w$.

Clearly, $w \in W$, $y = w + z$; moreover $z \in W^{\perp}$ since for each $u_j$ we have

$$\langle z, u_j \rangle = \langle y - w, u_j \rangle = \langle y, u_i \rangle - \langle w, u_j \rangle$$

$$= \langle y, u_i \rangle - \sum_{i=1}^{k} \langle y, u_i \rangle \underbrace{\langle u_i, u_j \rangle}_{=0 \text{ except when } i=j.}$$

$$= \langle y, u_i \rangle - \langle y, u_i \rangle = 0.$$

[Query: What remains? Uniqueness!]

Suppose $w' \in W$ and $z' \in W^\perp$ with $y = w' + z'$.

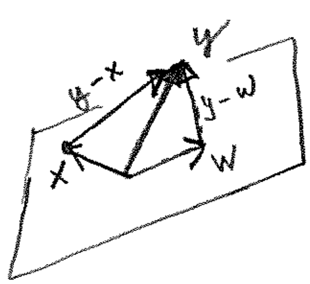
Then $w - w' = z' - z$ is in $W \cap W^\perp = \{0\}$

and so $w' = w$ and $z' = z$ as needed. $\blacksquare$

<u>Cor</u>: The vector $w = \text{proj}_W(y)$ above is the "closest" vector in $W$ to $y$ in the following sense:

$\|y - x\| \geq \|y - w\|$ for all $x \in W$, with equality only when $x = w$.

<u>Proof</u>: $\|y - x\|^2 = \|\overbrace{(w+z)}^{\text{as above}} - x\|^2 = \|\overbrace{(w-x)}^{\text{in } W} + \underbrace{z}_{\text{in } W^\perp}\|^2$



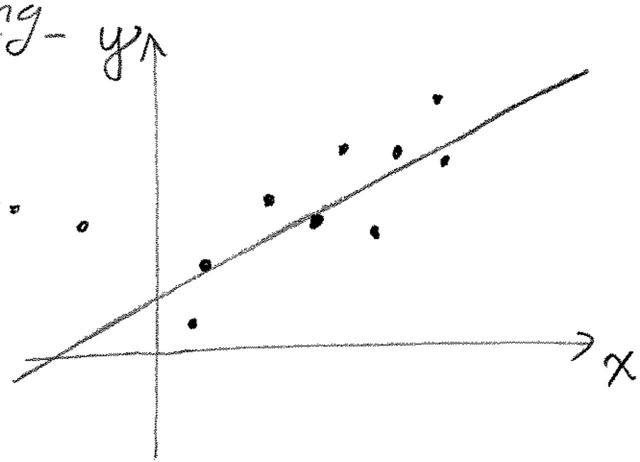

$$= \langle (w-x) + z, (w-x) + z \rangle$$

$$= \|w - x\|^2 + 0 + 0 + \|z\|^2$$

$$\geq \|z\|^2 = \|y - w\|^2$$

and can only have equality when $\|w - x\|^2 = 0 \Rightarrow w = x$. $\blacksquare$

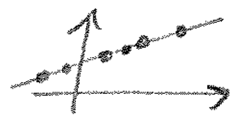
# Regression / Least Squares Fitting



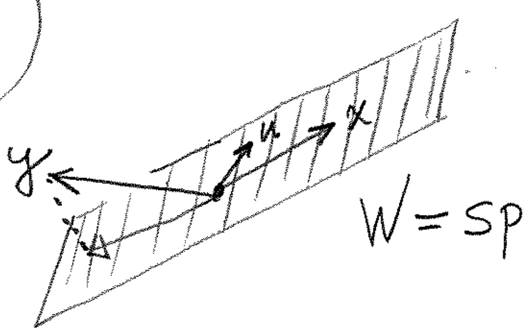

Data: $(x_i, y_i)$ for $i = 1, 2, \ldots, n$.

Which model $y = mx + b$ best fits this data?

In $\mathbb{R}^n$ consider
$$y = (y_1, \ldots, y_n)$$
$$x = (x_1, \ldots, x_n)$$
$$u = (1, \ldots, 1)$$

A perfect fit  corresponds to having $y \in \text{span}(\{x, u\})$. Pictorially, the general case is:

Space of data $= \mathbb{R}^n$:




$W = \text{span}(\{x, u\})$

Natural to define the best fit parameters $(m, b)$ to be the scalars with

$$\text{proj}_W(y) = mx + bu.$$

where here the projection is with respect to the usual dot product on $\mathbb{R}^n$. Concretely, this is the same as choosing $m, b$ to minimize

$$\sum_{i=1}^{n} \left(y_i - (mx_i - b)\right)^2$$

Note: Easily adapts to more complicated models.

Data: $(x_i, y_i, z_i)$ for $i = 1, 2, \ldots, n$

Model: $z = a x^2 + b x + c y + d \sin y$

Setup: In $\mathbb{R}^n$ consider $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$,
$z = (z_1, \ldots, z_n)$, $u = (x_1^2, \ldots, x_n^2)$, $v = (\sin y_1, \ldots, \sin y_n)$

Best Fit: $\mathrm{proj}_W(z)$ for $W = \mathrm{span}(\{u, x, y, v\})$.
is a linear combination $au + bx + cy + dv$.

[Q: How do we compute $\mathrm{proj}_W(z)$? ]

Thm: Suppose $\beta = \{w_1, \ldots, w_k\}$ is a basis for a subspace $W$ of $\mathbb{R}^n$. Let $A \in M_{k \times n}(\mathbb{R})$ be the matrix whose rows are $w_1, \ldots, w_k$.

Then $[\mathrm{proj}_W]^{\beta}_{\text{Std for } \mathbb{R}^n} = (A^t A)^{-1} A^t$

where $\mathrm{proj}_W : \mathbb{R}^n \to W$ is orthogonal projection with respect to the dot product on $\mathbb{R}^n$.

Proof: Some other time...