Def: Suppose $S$ is a nonempty subset of an inner product space $V$. The orthogonal complement 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)

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

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

Note: $S^\perp$ is always a subspace since if $x_1, x_2 \in S^\perp$ then $\langle cx_1 + x_2, y \rangle = c \langle x_1, y \rangle + \langle x_2, y \rangle = c \cdot 0 + 0 = 0$ for all $y \in S$.

Also $S \cap S^\perp$ contains at most the zero vector.
Thm: Suppose $W$ is a finite-dimensional 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 $\{u_1, u_2, \ldots, u_k\}$ 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

$$\text{proj}: V \rightarrow W.$$ 

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

$$\langle z, u_j \rangle = \langle y - w, u_j \rangle = \langle y, u_j \rangle - \langle w, u_j \rangle$$
\[ \langle y, u_i \rangle - \sum_{i=1}^{k} \langle y, u_i \rangle \langle u_i, u_j \rangle = 0 \] 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.

Cor: The vector \( w = \text{proj}_W(y) \) above is the "closest" vector in \( W \) to \( y \) in the following sense:
\[ \| y - x \| \geq \| y - w \| \text{ for all } x \in W, \text{ with equality only when } x = w. \]

Proof: \[ \| y - x \|^2 = \sqrt{\langle (w + z) - x, (w + z) - x \rangle} = \sqrt{\langle (w - x) + z, (w - x) + z \rangle} \]
\[ = \langle (w - x) + z, (w - x) + z \rangle \]
\[ = \| w - x \|^2 + 0 + 0 + \| z \|^2 \]
\[ \geq \| z \|^2 = \| y - w \|^2 \]
and can only have equality when \( \| w - x \|^2 = 0 \Rightarrow w = x. \)
Regression / Least Squares Fitting: $y$ vs. $x$

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

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

**Note:** Easily adapts to more complicated models.

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

**Model:** $z = ax^2 + bx + cy + d \sin y$

**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:** $\text{proj}_W(z)$ for $W = \text{span}(\{u, x, y, v\})$

is a linear combination $au + bx + cy + dv$.

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