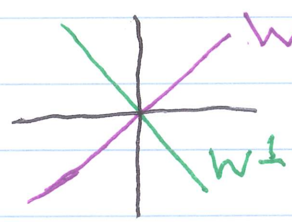
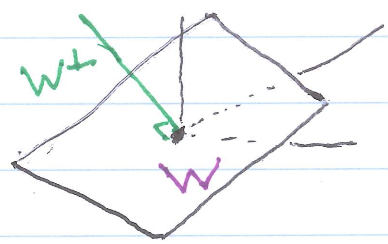
6.C Orthogonal complements and least squares

V = inner product space, $W \subseteq V$ subspace.
The orthogonal complement $W^\perp$ = all $x \in V$ such that $\langle x, w \rangle = 0$ for all $w \in W$.



$\mathbb{R}^2$



in $\mathbb{R}^3$

Simple facts.

$W^\perp$ is a subspace

$x, y \in W^\perp$ c scalar $\Longrightarrow$

$\langle x+y, w \rangle = \langle x, w \rangle + \langle y, w \rangle = 0 + 0 = 0$ all $w \in W$

$\langle cx, w \rangle = c \langle x, w \rangle = c \cdot 0 = 0$ all w.

$\{0\}^\perp = V$, $V^\perp = \{0\}$, $W \cap W^\perp = \{0\}$ all W

$W_1 \subseteq W_2 \Longrightarrow W_2^\perp \subseteq W_1^\perp$.

<u>Theorem</u> If V is finite-dimensional, then
$$W + W^{\perp} = V.$$

<u>Corollary</u> Every $x \in V$ has a unique decomp. $x_1 + x_2$ where $x_1 \in W$, $x_2 \in W^{\perp}$.

<u>Proof of Corollary</u> Existence of a decomposition is the conclusion of the theorem. For uniqueness, say $x_1 + x_2 = y_1 + y_2$ $\quad x_1, y_1 \in W$, $x_2, y_2 \in W^{\perp}$

Then $\underbrace{x_1 - y_1}_{\text{in } W} = \underbrace{y_2 - x_2}_{\text{in } W^{\perp}}$. Since $W \cap W^{\perp} = \{0\}$,

we must have $x_1 - y_1 = 0 = y_2 - x_2$, so that $x_1 = y_1$ and $x_2 = y_2$. ∎

<u>Proof of thm.</u> Let $\{w_1, \ldots, w_n\}$ be an orthonormal basis for W. If $x \in V$, then

$$x_2 = x - \underbrace{\sum_j \langle x, w_j \rangle w_j}_{= x_1} \in W^{\perp} \text{ by previous calculation.}$$

Thus $x_2 = x - x_1 \Rightarrow x = x_1 + x_2$, $x_1 \in W$, $x_2 \in W^{\perp}$. ∎

<u>Cor.</u> $\dim V = \dim W + \dim W^{\perp}$, or equivalently
$$\dim W^{\perp} = \dim V - \dim W.$$

<u>PROOF.</u> $W + W^{\perp} = V$, $W \cap W^{\perp} = \{0\} \Longrightarrow$

$$\dim W + \dim W^{\perp} = \dim(W + W^{\perp}) + \dim(W \cap W^{\perp})$$
$$= \dim V + 0. \blacksquare$$

<u>Corollary</u> $\dim V$ finite $\Longrightarrow$

$$W = (W^{\perp})^{\perp}.$$

<u>Proof.</u> $\dim W^{\perp\perp} = \dim W - \dim W^{\perp} = \dim W.$

On the other hand ~~$x \in (W^{\perp})^{\perp} \Rightarrow$~~

$x \in W \Longrightarrow \langle x, y \rangle = 0$ all $y \in W^{\perp} \Longrightarrow$ x belongs to $W^{\perp\perp}$.

Hence $W \subseteq W^{\perp\perp}$. If $r = \dim W = \dim W^{\perp\perp}$, then if $\{w_1, \ldots, w_r\}$ is a basis for W, it must also be a basis for $W^{\perp\perp}$, which has the same dimension. Hence $W \supseteq W^{\perp\perp}$, so $W = W^{\perp\perp}. \blacksquare$

Orthogonal Projection onto W

V = finite dimensional inner product space, $W \subseteq V$ subspace. The orthogonal (perpendicular) projection $E_W(v)$ of v onto W is defined by x_1, where $v = x_1 + x_2$ where $x_1 \in W$, $x_2 \in W^{\perp}$.

As before, if $\{w_1, \dots, w_r\}$ is an orthonormal basis for W, then $E_W(x) = \sum \langle x, w_j \rangle w_j$.

<u>Prop (~~missing from~~ obscure in Axler!)</u> E_W is linear.

<u>Verification</u> $E_W(x+y) = \sum \langle x+y, w_j \rangle w_j = \dots$

$$\sum \langle x, w_j \rangle w_j + \langle y, w_j \rangle w_j = \dots = E_W(x) + E_W(y).$$

$$E_W(cx) = \sum \langle cx, w_j \rangle w_j = \sum c \langle x, w_j \rangle w_j =$$

$$c \sum c \langle x, w_j \rangle w_j = c\, E_W(x). \blacksquare$$

<u>Further properties</u>

(i) $E_W(x) = x$ if $x \in W$, $E_W(x) = 0$ if $x \in W^\perp$.

(ii) $(E_W)^2 x = E_W x$.

<u>Verification of (ii)</u> Write $x = x_1 + x_2$ as before. Then $E^2 x = E(Ex) = E(E(x_1 + x_2)) =$ $EEx_1 = Ex_1 = x_1$. But $x_1 = Ex$. Since $E^2 x = Ex$ for all x, we have $E^2 = E$. $\blacksquare$

Example Let $W =$ Span of $(2,3,4)$ and $(0,3,4)$. If $x = (1,1,0)$, find $E_W(x)$.

Sketch of computation First find orthonormal basis for W. Gram-Schmidt yields $(0,\frac{3}{5},\frac{4}{5})$ & $(1,0,0)$. Therefore $E_W(1,1,0) =$

$$\langle (0,\tfrac{3}{5},\tfrac{4}{5}),(1,1,0)\rangle (0,\tfrac{3}{5},\tfrac{4}{5}) + \langle (1,0,0),(1,1,0)\rangle (1,0,0) =$$

$$\frac{3}{5}\cdot\left(0,\frac{3}{5},\frac{4}{5}\right) + 1\cdot(1,0,0) = \left(1,\frac{9}{25},\frac{12}{25}\right).$$

Least squares principle

$W \subseteq V$ fin dim innerproduct space.

For $x \in V$, find $y \in W$ so that $|x-y|^2$ is as small as possible.

Solution The minimum occurs when $y = E_W x$ and nowhere else.

~~Verification Write $y = y_1 + y_2$ where $y_1 \in$~~

Verification Write $x = x_1 + x_2$ where $x_1 \in W$, $x_2 \in W^\perp$. Then $|(x-y)|^2 =$ $|x_2 + (x_1 - y)|^2$, and since $x_2 \perp x_1 - y$ we know that ↑ equals $|x_2|^2 + |x_1 - y|^2 \geq |x_2|^2$. Equality holds $\Leftrightarrow$ $x_1 - y = 0$ or $y = x_1 = E_W x$. ∎

Cor. If $u_1, \ldots, u_r$ is an orthonormal basis for W, then $|x - \sum c_j u_j|^2$ is minimized $\Leftrightarrow$

$c_j = \langle x, u_j \rangle$ all j. ∎

Fundamental Application: Want an empirical formula for one variable y as $\sum_{j=1}^{n} a_j x_j + b$ in terms of variables $x_1, \ldots, x_n$. We may try to do this by measuring $x_1, \ldots, x_n, y$ in m cases, where m is much larger than $n+1$. Usually it is not possible to find an exact solution for $a_1, \ldots, a_n, b$. The best we can hope for is a "least squares soln."

Suppose we are given column vectors

$X_1, \ldots, X_n \qquad X_j = \begin{pmatrix} x_{1j} \\ \vdots \\ x_{mj} \end{pmatrix}$ of m observations

in which the i th observation yields x_{ij} for variable x_j and similarly for Y. Let $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = U$. Find constants $a_1, \ldots, a_n, b$ so that

$$\left| Y - \sum a_j X_j + bU \right|^2$$ is minimized.

If $\{X_1, \ldots, X_n, U\}$ is linearly independant the methods of this section and the previous one can be used to find the coefficients $a_1, \ldots, a_n, b$. However, in most cases it is much better to use more systematic methods that would require too much time to describe (but are not much more complicated!).