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

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

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

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

<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 $E^2 = E$. ∎

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~~ [illegible]

<u>Verification</u> 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 $\updownarrow$ equals $|x_2|^2 + |x_1 - y|^2 \geq |x_2|^2$.
Equality holds $\Longleftrightarrow$ $x_1 - y = 0$ or $y = x_1 = E_W x$. ∎

<u>Cor.</u> If $u_1, \dots, u_r$ is an orthonormal basis for W, then $|x - \sum c_j u_j|^2$ is minimized $\Longleftrightarrow$

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

<u>Fundamental Application</u>: Want an empirical formula for one variable y as $\sum_{j=1}^{n} a_j x_j + b$ in terms of variables $x_1, \dots, x_n$.

We may try to do this by measuring $x_1, \dots, x_n, y$ in m cases, where m is much larger than $n+1$. Usually it is not possible to find an exact solution for $a_1, \dots, a_n, b$. The best we can hope for is a "least squares soln."