Review: Orthogonal Decomposition Theorem

W a subspace of $\mathbb{R}^n$. Each $\boldsymbol{y} \in \mathbb{R}^n$ can be written in a unique way as

$$\boldsymbol{y} = \widehat{\boldsymbol{y}} + \boldsymbol{z} \text{ where } \begin{cases} \widehat{\boldsymbol{y}} & \text{is in } W \\ \boldsymbol{z} & \text{is in } W^{\perp} \end{cases}$$

If $W = \{\boldsymbol{u_1}, \ldots, \boldsymbol{u_p}\}$ is an orthogonal basis for W, then formula are

$$\begin{aligned} \widehat{\boldsymbol{y}} &= \tfrac{\boldsymbol{y}\cdot\boldsymbol{u_1}}{\boldsymbol{u_1}\cdot\boldsymbol{u_1}}\boldsymbol{u_1} + \ldots + \tfrac{\boldsymbol{y}\cdot\boldsymbol{u_p}}{\boldsymbol{u_p}\cdot\boldsymbol{u_p}}\boldsymbol{u_p} \\ &\quad \updownarrow \qquad\qquad \updownarrow \\ &= \text{proj}_{\boldsymbol{u_1}}\boldsymbol{y} + \ldots + \text{proj}_{\boldsymbol{u_p}}\boldsymbol{y} \\ &= \text{proj}_{W}\boldsymbol{y} \end{aligned}$$

$\widehat{\boldsymbol{y}}$ is called the orthogonal projection of y

and $\boldsymbol{z} = \boldsymbol{y} - \widehat{\boldsymbol{y}}$

$\boldsymbol{z}$ is called the component of $\boldsymbol{y}$ orthogonal to W

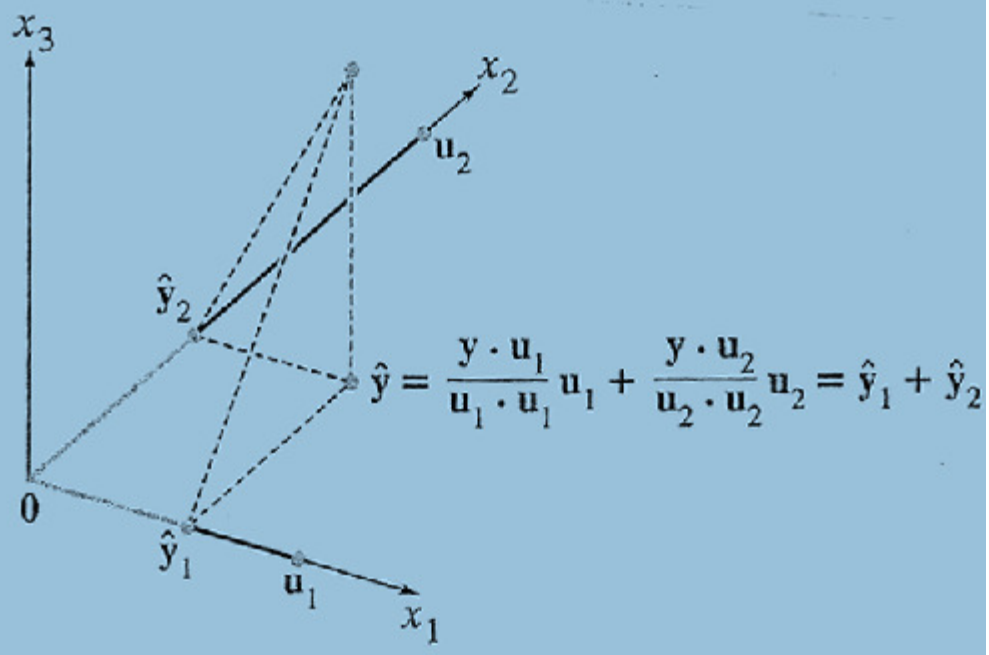


FIGURE 3 The orthogonal projection of $\mathbf{y}$ is the sum of its projections onto one-dimensional subspaces that are mutually orthogonal.

$$\begin{aligned} \widehat{\boldsymbol{y}} &= \frac{\boldsymbol{y}\cdot\boldsymbol{u}_1}{\boldsymbol{u}_1\cdot\boldsymbol{u}_1}\boldsymbol{u}_1 + \frac{\boldsymbol{y}\cdot\boldsymbol{u}_p}{\boldsymbol{u}_p\cdot\boldsymbol{u}_p}\boldsymbol{u}_2 \\ &\quad\updownarrow \qquad\qquad \updownarrow \\ &= \text{proj}_{u_1}\boldsymbol{y} + \text{proj}_{u_2}\boldsymbol{y} \end{aligned}$$

If $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \ \ldots, \boldsymbol{u_p}\}$ is an orthonormal basis for the subspace W, then "projection onto W" can be accomplished by a matrix multiplication.

Let $U = [\boldsymbol{u_1}\ \boldsymbol{u_2}\ \ \ldots\ \ \boldsymbol{u_p}]$ (so that $W = \operatorname{col} U$).

Since U has orthonormal columns, we know that $U^T U = I$.

It turns out (*see textbook/lecture notes*) that for any $\boldsymbol{y}$ in $\mathbb{R}^n$

$$UU^T\boldsymbol{y} = \widehat{\boldsymbol{y}} = \mathbf{proj}_W \boldsymbol{y} \text{ where } W = \operatorname{col} U.$$

Example Let $W = \text{Span}\{\begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}2\\0\\1\\1\end{bmatrix}, \begin{bmatrix}0\\0\\-1\\1\end{bmatrix}\} \subseteq \mathbb{R}^4$ and $\boldsymbol{y} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix} \notin W$

$$\widehat{\boldsymbol{y}} = \tfrac{\boldsymbol{y}\cdot\boldsymbol{u_1}}{\boldsymbol{u_1}\cdot\boldsymbol{u_1}}\boldsymbol{u_1} + \tfrac{\boldsymbol{y}\cdot\boldsymbol{u_2}}{\boldsymbol{u_2}\cdot\boldsymbol{u_2}}\boldsymbol{u_2} + \tfrac{\boldsymbol{y}\cdot\boldsymbol{u_3}}{\boldsymbol{u_3}\cdot\boldsymbol{u_3}}\boldsymbol{u_3}$$

$$= \tfrac{1}{1}\begin{bmatrix}0\\1\\0\\0\end{bmatrix} + \tfrac{4}{6}\begin{bmatrix}\mathbf{2}\\\mathbf{0}\\\mathbf{1}\\\mathbf{1}\end{bmatrix} + \tfrac{0}{2}\begin{bmatrix}0\\0\\-1\\1\end{bmatrix} = \begin{bmatrix}\frac{4}{3}\\1\\\frac{2}{3}\\\frac{2}{3}\end{bmatrix} = \text{point in } W \text{ closest to } \boldsymbol{y}$$

Distance from $\boldsymbol{y}$ to $W = ||\boldsymbol{z}|| = ||\,\boldsymbol{y} - \widehat{\boldsymbol{y}}\,|| = \frac{\sqrt{3}}{3}$

If we like, we could normalize the basis to get an orthonormal basis for W :

$$\left\{\begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}2/\sqrt{6}\\0\\1/\sqrt{6}\\1/\sqrt{6}\end{bmatrix}, \begin{bmatrix}0\\0\\-1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}\right\}$$

If we let $U = \begin{bmatrix}0 & 2/\sqrt{6} & 0\\1 & 0 & 0\\0 & 1/\sqrt{6} & -1/\sqrt{2}\\0 & 1/\sqrt{6} & 1/\sqrt{2}\end{bmatrix}_{4\times 3}$ (so $W = \operatorname{col} U$)

$$U^T = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 2/\sqrt{6} & 0 & 1/\sqrt{6} & 1/\sqrt{6} \\ 0 & 0 & -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}_{3\times 4}$$

and $UU^T = \begin{bmatrix} \frac{2}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & 0 & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}_{4\times 4}$

$T: \mathbb{R}^4 \to \mathbb{R}^4, \quad T(\boldsymbol{y}) = UU^T\boldsymbol{y} = \widehat{\boldsymbol{y}}$

is the orthogonal projection : $\mathbb{R}^4$ onto $W = \operatorname{Col} U$:

So (using the same $\boldsymbol{y} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ as above):

$$UU^T\boldsymbol{y} = UU^T \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & 0 & -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{4}{3} \\ 1 \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix} = \widehat{\boldsymbol{y}} \quad \text{(same answer as before)}$$

Gram Schmidt Formulas

(*see notes posted in syllabus about the Gram Schmidt Process*)

Suppose $\{\boldsymbol{x_1}, ..., \boldsymbol{x_p}\}$ is a basis for the subspace W of $\mathbb{R}^n$. Then we can convert to an orthogonal basis $\{\boldsymbol{v_1}, \boldsymbol{v_2}, ..., \boldsymbol{v_p}\}$ is an orthogonal basis for W using the formulas:

$$\boldsymbol{v_1} = \boldsymbol{x_1}$$

$$\begin{aligned}\boldsymbol{v_2} &= \boldsymbol{x_2} - \text{proj}_{\boldsymbol{v_1}}\boldsymbol{x_2} \\ &= \boldsymbol{x_2} - \frac{\boldsymbol{x_2}\cdot\boldsymbol{v_1}}{\boldsymbol{v_1}\cdot\boldsymbol{v_1}}\boldsymbol{v_1}\end{aligned}$$

$$\begin{aligned}\boldsymbol{v_3} &= \boldsymbol{x_3} - (\text{proj}_{\boldsymbol{v_1}}\boldsymbol{x_3} + \text{proj}_{\boldsymbol{v_2}}\boldsymbol{x_3}) \\ &= \boldsymbol{x_3} - \frac{\boldsymbol{x_3}\cdot\boldsymbol{v_1}}{\boldsymbol{v_1}\cdot\boldsymbol{v_1}}\boldsymbol{v_1} - \frac{\boldsymbol{x_3}\cdot\boldsymbol{v_2}}{\boldsymbol{v_2}\cdot\boldsymbol{v_2}}\boldsymbol{v_2}\end{aligned}$$

$$\vdots$$

$$\begin{aligned}\boldsymbol{v_p} &= \boldsymbol{x_p} - (\text{proj}_{\boldsymbol{v_1}}\boldsymbol{x_p} + \text{proj}_{\boldsymbol{v_2}}\boldsymbol{x_p} - ... \quad ... + \text{proj}_{\boldsymbol{v_{p-1}}}\boldsymbol{x_p} \\ &= \boldsymbol{x_p} - \frac{\boldsymbol{x_p}\cdot\boldsymbol{v_1}}{\boldsymbol{v_1}\cdot\boldsymbol{v_1}}\boldsymbol{v_1} - \frac{\boldsymbol{x_p}\cdot\boldsymbol{v_2}}{\boldsymbol{v_2}\cdot\boldsymbol{v_2}}\boldsymbol{v_2} - ... - \frac{\boldsymbol{x_p}\cdot\boldsymbol{v_{p-1}}}{\boldsymbol{v_{p-1}}\cdot\boldsymbol{v_{p-1}}}\boldsymbol{v_{p-1}}\end{aligned}$$

Example: $W = \text{Span}\left\{\begin{bmatrix}0\\2\\0\\1\end{bmatrix}, \begin{bmatrix}1\\0\\0\\1\end{bmatrix}, \begin{bmatrix}1\\0\\1\\1\end{bmatrix}\right\}$, a basis for W

$\uparrow \quad \uparrow \quad \uparrow$

$\boldsymbol{x_1} \quad \boldsymbol{x_2} \quad \boldsymbol{x_3}$

Find an orthogonal basis for W

Let $\boldsymbol{v_1} = \boldsymbol{x_1} = \begin{bmatrix}0\\2\\0\\1\end{bmatrix}$

$W_1 = \text{Span}\{\boldsymbol{x_1}\} = \text{Span}\{\boldsymbol{v_1}\}$.

Let $W_2 = \text{Span}\{\boldsymbol{x_1}, \boldsymbol{x_2}\} = \text{Span}\{\boldsymbol{v_1}, \boldsymbol{x_2}\}$.

Replace $\boldsymbol{x_2}$ with $\boldsymbol{v_2}$ = the component of $\boldsymbol{x_2}$ orthogonal to W_1

$$= \boldsymbol{x_2} - \text{proj}_{\boldsymbol{v_1}}\boldsymbol{x_2}$$

$$\text{proj}_{\boldsymbol{v_1}} \boldsymbol{x_2} = \frac{\boldsymbol{x_2}\cdot\boldsymbol{v_1}}{\boldsymbol{v_1}\cdot\boldsymbol{v_1}}\,\boldsymbol{v_1}$$

$$= \frac{\begin{bmatrix}1\\0\\0\\1\end{bmatrix}\cdot\begin{bmatrix}0\\2\\0\\1\end{bmatrix}}{\begin{bmatrix}0\\2\\0\\1\end{bmatrix}\cdot\begin{bmatrix}0\\2\\0\\1\end{bmatrix}}\begin{bmatrix}0\\2\\0\\1\end{bmatrix} = \frac{1}{5}\begin{bmatrix}0\\2\\0\\1\end{bmatrix} = \begin{bmatrix}0\\\frac{2}{5}\\0\\\frac{1}{5}\end{bmatrix}, \quad \text{so}$$

$$\boldsymbol{v_2} = \boldsymbol{x_2} - \text{proj}_{\boldsymbol{v_2}}\boldsymbol{x_2} = \begin{bmatrix}1\\0\\0\\1\end{bmatrix} - \begin{bmatrix}0\\\frac{2}{5}\\0\\\frac{1}{5}\end{bmatrix} = \begin{bmatrix}1\\-\frac{2}{5}\\0\\\frac{4}{5}\end{bmatrix}$$

Optional: rescale $\boldsymbol{v_2}$ to avoid working with the fractions in the remaining calculations:

$$\text{new } \boldsymbol{v_2} = 5\begin{bmatrix}1\\-\frac{2}{5}\\0\\\frac{4}{5}\end{bmatrix} = \begin{bmatrix}5\\-2\\0\\4\end{bmatrix}$$

Then $\text{Span}\{\boldsymbol{x_1}, \boldsymbol{x_2}\} = \text{Span}\{\boldsymbol{v_1}, \boldsymbol{x_2}\} = \text{Span}\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$.

Let $W_3 = \text{Span}\{\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}\} = \text{Span}\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{x_3}\}$
$\quad\;\|$
$\quad W$

Replace $\boldsymbol{x_3}$ with $\boldsymbol{v_3}$ = component of $\boldsymbol{x_3}$ orthog. to W_2

$$= \boldsymbol{x_3} - \text{proj}_{W_2}\boldsymbol{x_3}$$
$$= \boldsymbol{x_3} - (\text{proj}_{\boldsymbol{v_1}}\boldsymbol{x_3} + \text{proj}_{\boldsymbol{v_2}}\boldsymbol{x_3})$$

$$\text{proj}_{W_2}\boldsymbol{x_3} = \frac{\boldsymbol{x_3}\cdot\boldsymbol{v_1}}{\boldsymbol{v_1}\cdot\boldsymbol{v_1}}\,\boldsymbol{v_1} + \frac{\boldsymbol{x_3}\cdot\boldsymbol{v_2}}{\boldsymbol{v_2}\cdot\boldsymbol{v_2}}$$

$$= \frac{\begin{bmatrix}1\\0\\1\\1\end{bmatrix}\cdot\begin{bmatrix}0\\2\\0\\1\end{bmatrix}}{\begin{bmatrix}0\\2\\0\\1\end{bmatrix}\cdot\begin{bmatrix}0\\2\\0\\1\end{bmatrix}}\begin{bmatrix}0\\2\\0\\1\end{bmatrix} + \frac{\begin{bmatrix}1\\0\\1\\1\end{bmatrix}\cdot\begin{bmatrix}5\\-2\\0\\4\end{bmatrix}}{\begin{bmatrix}5\\-2\\0\\4\end{bmatrix}\cdot\begin{bmatrix}5\\-2\\0\\4\end{bmatrix}}\begin{bmatrix}5\\-2\\0\\4\end{bmatrix} = \quad \cdots \quad = \begin{bmatrix}1\\0\\0\\1\end{bmatrix},$$

so $\boldsymbol{v_3} = \boldsymbol{x_3} - \text{proj}_{W_2}\boldsymbol{x_3} = \begin{bmatrix}1\\0\\1\\1\end{bmatrix} - \begin{bmatrix}1\\0\\0\\1\end{bmatrix} = \begin{bmatrix}0\\0\\1\\0\end{bmatrix}$

Then $W_3 = \text{Span}\{\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}\} = \text{Span}\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$, and we have replaced all the $\boldsymbol{x_i}$'s. W_3 is the

An orthogonal basis for $W_3 = W$ is

$$\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\} = \left\{\begin{bmatrix}0\\2\\0\\1\end{bmatrix}, \begin{bmatrix}5\\-2\\0\\4\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix}\right\}$$

At the start: you can check that $\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}$ are linearly independent and therefore form a basis for W.

What would happen if we applied the Gram-Schmidt Process to a linearly dependent set of vectors $\{\boldsymbol{x_1}, ..., \boldsymbol{x_p}\}$?