Most important ideas:

- For a given subspace W of R^n, every vector $\vec{y}$ in R^n can be written as the sum of a vector $\hat{\vec{y}}$ in W and a vector $\vec{z} = \vec{y} - \hat{\vec{y}}$ orthogonal to W.
- The projection $\hat{\vec{y}}$ of vector $\vec{y}$ onto a subspace W is the part of $\vec{y}$ that is in W.
- How to find the projection of $\vec{y}$ onto W if we have an orthogonal basis for W?
- What it means for one vector space to be orthogonal to another vector space: every vector in the one vector is orthogonal to every vector in the other.

We saw in 6.1 that every subspace W of R^n has an orthogonal complement $W^\perp$, which is the collection of all vectors that are orthogonal to all of the vectors in W.

Example 1, in R^3: **(1) if W is (all of the vectors in) a plane through the origin, then $W^\perp$ is (all of the vectors in) the line through the origin that is perpendicular to that plane; (2) if W is (all of the vectors on) a line through the origin, then $W^\perp$ is (all of the vectors in) the plane through the origin that is perpendicular to that line. This is like Figure 7 on page 334.**

Example 2, in R^4: Let $W = span\left\{\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\0\\1\end{bmatrix}\right\}$. Call these $\vec{u}_1, \vec{u}_2$. Find $W^\perp$, that is, find the collection of all vectors $\vec{x} = \begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}$ in R^4 such that $\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} \cdot \begin{bmatrix}1\\-1\\-1\\1\end{bmatrix} = 0$ and $\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} \cdot \begin{bmatrix}0\\1\\0\\1\end{bmatrix} = 0$.

That is, $\begin{bmatrix}1 & -1 & -1 & 1\\0 & 1 & 0 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$ which leads to $\left[\begin{array}{cccc|c}1 & -1 & -1 & 1 & 0\\0 & 1 & 0 & 1 & 0\end{array}\right] \sim \left[\begin{array}{cccc|c}1 & 0 & -1 & 2 & 0\\0 & 1 & 0 & 1 & 0\end{array}\right]$.

So $\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} = \begin{bmatrix}x_3 - 2x_4\\-x_4\\x_3\\x_4\end{bmatrix} = x_3\begin{bmatrix}1\\0\\1\\0\end{bmatrix} + x_4\begin{bmatrix}-2\\-1\\0\\1\end{bmatrix}$, so $W^\perp = span\left\{\begin{bmatrix}1\\0\\1\\0\end{bmatrix}, \begin{bmatrix}-2\\-1\\0\\1\end{bmatrix}\right\}$. Call these $\vec{u}_3$ and $\vec{u}_4$.

Notice that $\vec{u}_1 \cdot \vec{u}_3 = \begin{bmatrix}1\\-1\\-1\\1\end{bmatrix} \cdot \begin{bmatrix}1\\0\\1\\0\end{bmatrix} = 0$, and similarly, $\vec{u}_1 \cdot \vec{u}_4 = 0, \vec{u}_2 \cdot \vec{u}_3 = 0, \vec{u}_2 \cdot \vec{u}_4 = 0$.

So any vector $c_1\vec{u}_1 + c_2\vec{u}_2$ in W is orthogonal to any vector $c_3\vec{u}_3 + c_4\vec{u}_4$ in $W^\perp$:

$$\begin{aligned}(c_1\vec{u}_1 + c_2\vec{u}_2) \cdot (c_3\vec{u}_3 + c_4\vec{u}_4) &= c_1c_3\,\vec{u}_1 \cdot \vec{u}_3 + c_1c_4\,\vec{u}_1 \cdot \vec{u}_4 + c_2c_3\,\vec{u}_2 \cdot \vec{u}_3 + c_2c_4\,\vec{u}_2 \cdot \vec{u}_4\\ &= c_1c_3(0) + c_1c_4(0) + c_2c_3(0) + c_2c_4(0) = 0\end{aligned}$$

We say that vector space W (which is a subspace of R^4) is orthogonal to the vector space $W^\perp$ (another subspace of R^4). Too bad we can't visualize any of this in 4 dimensions.

It turn out this is useful: Any vector in R^4 can be split into two parts: one part that comes from W and the other part that comes from $W^\perp$. Back to this thought in a bit.

Recall Theorem 5 from Section 6.2:

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each $\mathbf{y}$ in W, the weights in the linear combination

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

are given by

$$c_j = \frac{\mathbf{y}\cdot\mathbf{u}_j}{\mathbf{u}_j\cdot\mathbf{u}_j} \qquad (j = 1, \ldots, p)$$

Theorem 8 in Section 6.3 on page 348 says something similar:

Let W be a subspace of $\mathbb{R}^n$. Then each $\mathbf{y}$ in $\mathbb{R}^n$ can be written uniquely in the form

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \qquad (1)$$

where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$. In fact, if $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is any orthogonal basis of W, then

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}\mathbf{u}_p \qquad (2)$$

and $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

What's the difference? Now we are considering <u>any</u> vector $\vec{y}$ in R^n, not just vectors <u>in</u> W. So

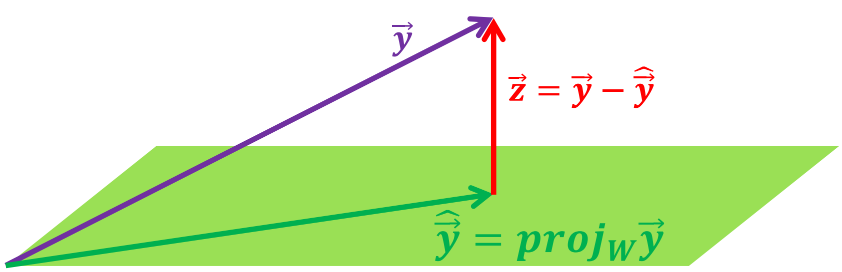
$$\vec{y} = \hat{\vec{y}} + \vec{z}$$

where

$$\hat{\vec{y}} = proj_W\vec{y} \text{ is the part of } \vec{y} \text{ in } W$$

and

$$\vec{z} = \vec{y} - \hat{\vec{y}} \text{ is the } \textbf{part} \text{ of } \vec{y} \text{ orthogonal to } W$$
$$\text{that is, the } \textbf{part} \text{ of } \vec{y} \text{ in } W^\perp$$



$$W = span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\}$$

where (assuming $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p$ are mutually orthogonal)

$$\hat{\vec{y}} = \frac{\vec{y}\cdot\vec{u}_1}{\vec{u}_1\cdot\vec{u}_1}\vec{u}_1 + \frac{\vec{y}\cdot\vec{u}_2}{\vec{u}_2\cdot\vec{u}_2}\vec{u}_2 + \cdots + \frac{\vec{y}\cdot\vec{u}_p}{\vec{u}_p\cdot\vec{u}_p}\vec{u}_p\,.$$

Example 3, in R^4: in R^4. $\vec{y} = \begin{bmatrix} 11 \\ -1 \\ -1 \\ -5 \end{bmatrix}$, $\vec{u}_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$, $\vec{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$, and $W = span\{\vec{u}_1, \vec{u}_2\}$.

Then

$$\hat{\vec{y}} = proj_W \vec{y} = \frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} + \frac{\vec{y} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \frac{8}{4} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} + \frac{-6}{2} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -5 \\ -2 \\ -1 \end{bmatrix}$$

and $\vec{z} = \vec{y} - \hat{\vec{y}} = \begin{bmatrix} 11 \\ -1 \\ -1 \\ -5 \end{bmatrix} - \begin{bmatrix} 2 \\ -5 \\ -2 \\ -1 \end{bmatrix} = \begin{bmatrix} 9 \\ 4 \\ 1 \\ -4 \end{bmatrix}$, which is $\perp$ $\vec{u}_1$ and $\vec{u}_2$, $\begin{bmatrix} 9 \\ 4 \\ 1 \\ -4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} = 0$, $\begin{bmatrix} 9 \\ 4 \\ 1 \\ -4 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} = 0$,

and thus is in $W^{\perp}$.

For $W = span\left\{\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}\right\}$, on Page 1 of this handout we found $W^{\perp} = span\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \end{bmatrix}\right\}$,

and (after a little work) we could find that $\begin{bmatrix} 9 \\ 4 \\ 1 \\ -4 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - 4 \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \end{bmatrix}$.

So $\vec{y}$ is the sum of **a vector in W** and **a vector in $W^{\perp}$**:

$$\vec{y} = \begin{bmatrix} 11 \\ -1 \\ -1 \\ -5 \end{bmatrix} = \begin{bmatrix} 2 \\ -5 \\ -2 \\ -1 \end{bmatrix} + \begin{bmatrix} 9 \\ 4 \\ 1 \\ -4 \end{bmatrix}.$$

The part of $\vec{y}$ in W ↑ **↑ The part of $\vec{y}$ in $W^{\perp}$**

One final interesting thought—see Theorem 10, page 351. Suppose $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\}$ is an orthonormal basis for W, then

$$\begin{aligned}
\hat{\vec{y}} &= proj_W \vec{y} \\
&= \frac{\vec{y}\cdot\vec{u}_1}{\vec{u}_1\cdot\vec{u}_1}\,\vec{u}_1 + \frac{\vec{y}\cdot\vec{u}_2}{\vec{u}_2\cdot\vec{u}_2}\,\vec{u}_2 + \cdots + \frac{\vec{y}\cdot\vec{u}_p}{\vec{u}_p\cdot\vec{u}_p}\,\vec{u}_p \quad \text{since } \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\} \text{ is orthogonal} \\
&= \vec{y}\cdot\vec{u}_1\ \ \vec{u}_1 + \vec{y}\cdot\vec{u}_2\ \ \vec{u}_2 + \cdots + \vec{y}\cdot\vec{u}_p\ \ \vec{u}_p \quad \text{since } \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\} \text{ is orthonormal} \\
&= (\vec{u}_1^T\vec{y})\,\vec{u}_1 + (\vec{u}_2^T\vec{y})\,\vec{u}_2 + \cdots + (\vec{u}_p^T\vec{y})\,\vec{u}_p \\
&= [\,\vec{u}_1\ \vec{u}_2\ \cdots\ \vec{u}_p\,]\begin{bmatrix}\vec{u}_1^T\vec{y}\\ \vec{u}_2^T\vec{y}\\ \vdots\\ \vec{u}_p^T\vec{y}\end{bmatrix} \\
&= UU^T\vec{y} \text{ where } U = [\vec{u}_1\ \vec{u}_2 \cdots\ \vec{u}_p].
\end{aligned}$$

Whoa! So the projection of $\vec{y}$ onto $W = span\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\}$ is $\hat{\vec{y}} = proj_W\vec{y} = UU^T\vec{y}$.

In the previous example:

$$U = \begin{bmatrix}1 & 0\\ -1 & 1\\ -1 & 0\\ 1 & 1\end{bmatrix} \text{ so } UU^T = \begin{bmatrix}1 & 0\\ -1 & 1\\ -1 & 0\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & -1 & -1 & 1\\ 0 & 1 & 0 & 1\end{bmatrix} = \begin{bmatrix}1/4 & -1/4 & -1/4 & 1/4\\ -1/4 & 3/4 & 1/4 & 1/4\\ -1/4 & 1/4 & 1/4 & -1/4\\ 1/4 & 1/4 & -1/4 & 3/4\end{bmatrix}.$$

So $\hat{\vec{y}} = proj_W\vec{y} = UU^T\vec{y} = \begin{bmatrix}1/4 & -1/4 & -1/4 & 1/4\\ -1/4 & 3/4 & 1/4 & 1/4\\ -1/4 & 1/4 & 1/4 & -1/4\\ 1/4 & 1/4 & -1/4 & 3/4\end{bmatrix}\begin{bmatrix}11\\ -1\\ -1\\ -5\end{bmatrix} = \begin{bmatrix}2\\ -5\\ -2\\ -1\end{bmatrix}.$

Note: UU^T is always symmetric since $(UU^T)^T = (U^T)^TU^T = UU^T$.

Note: **if $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ is a basis for R^n, then** $U = [\vec{u}_1\ \vec{u}_2 \cdots\ \vec{u}_n]$ is square, and thus $U^TU = I \Rightarrow UU^T = I$ which means that $\hat{\vec{y}} = proj_W\vec{y} = UU^T\vec{y} = I\vec{y} = \vec{y}$. That is, we can exactly (rather than approximately) build $\vec{y}$ as a linear combination of $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$, as described in Theorem 5 in Section 6.2.

Summary: if U is orthogonal (its columns are orthonormal), then $U^TU = I$ and UU^T is used to find $\hat{\vec{y}} = proj_W\vec{y} = UU^T\vec{y}$, which is the best approximation of $\vec{y}$ (that is, $\hat{\vec{y}}$ is the vector closest to $\vec{y}$) that can be built from (as a linear combination of) the columns of U.