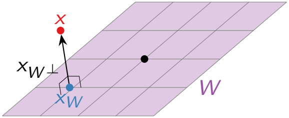
Section 6.3

Orthogonal Projections

Idea Behind Orthogonal Projections

If x is not in a subspace W, then y in W is the closest to x if $x - y$ is in $W^{\perp}$:



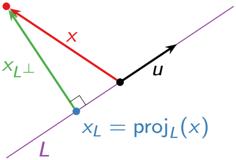
Reformulation: Every vector x can be decompsed uniquely as

$$x = x_W + x_{W^\perp}$$

where $x_W = y$ is the closest vector to x in W, and $x_{W^\perp} = x - y$ is in $W^{\perp}$.

Example: Let $u = \binom{3}{2}$ and let $L = \mathrm{Span}\{u\}$. Let $x = \binom{-6}{4}$. Then the closest point to x in L is $\mathrm{proj}_L(x) = \frac{x \cdot u}{u \cdot u} u$, so

$$x_L = \mathrm{proj}_L(x) = -\frac{10}{13}\begin{pmatrix}3\\2\end{pmatrix} \qquad x_{L^\perp} = x - \mathrm{proj}_L(x) = \begin{pmatrix}-6\\4\end{pmatrix} + \frac{10}{13}\begin{pmatrix}3\\2\end{pmatrix}.$$



Orthogonal Projections

Definition

Let W be a subspace of $\mathbf{R}^n$, and let $\{u_1, u_2, \ldots, u_m\}$ be an *orthogonal* basis for W. The **orthogonal projection** of a vector x onto W is

$$\mathrm{proj}_W(x) \stackrel{\mathrm{def}}{=} \sum_{i=1}^{m} \frac{x \cdot u_i}{u_i \cdot u_i}\, u_i.$$

Question: What is the difference between this and the formula for $[x]_{\mathcal{B}}$ from before?

Theorem

Let W be a subspace of $\mathbf{R}^n$, and let x be a vector in $\mathbf{R}^n$. Then $\mathrm{proj}_W(x)$ is the closest point to x in W. Therefore

$$\boxed{x_W = \mathrm{proj}_W(x) \qquad x_{W^\perp} = x - \mathrm{proj}_W(x).}$$

Why? Let $y = \mathrm{proj}_W(x)$. We need to show that $x - y$ is in $W^\perp$. In other words, $u_i \cdot (x - y) = 0$ for each i. Let's do u_1:

$$u_1 \cdot (x - y) = u_1 \cdot \left(x - \sum_{i=1}^{m} \frac{x \cdot u_i}{u_i \cdot u_i}\, u_i \right) = u_1 \cdot x - \frac{x \cdot u_1}{u_1 \cdot u_1}(u_1 \cdot u_1) - 0 - \cdots = 0.$$

Orthogonal Projections

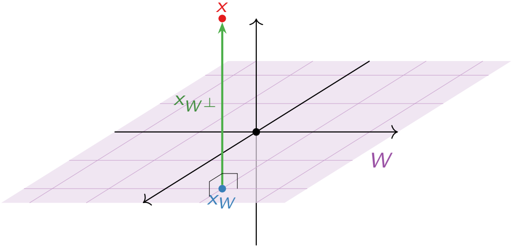
Easy example

What is the projection of $x = \begin{pmatrix}1\\2\\3\end{pmatrix}$ onto the xy-plane?

Answer: The xy-plane is $W = \mathrm{Span}\{e_1, e_2\}$, and $\{e_1, e_2\}$ is an orthogonal basis.

$$x_W = \mathrm{proj}_W \begin{pmatrix}1\\2\\3\end{pmatrix} = \frac{x \cdot e_1}{e_1 \cdot e_1} e_1 + \frac{x \cdot e_2}{e_2 \cdot e_2} e_2 = \frac{1 \cdot 1}{1^2} e_1 + \frac{1 \cdot 2}{1^2} e_2 = \begin{pmatrix}1\\2\\0\end{pmatrix}.$$

So this is the same projection as before.



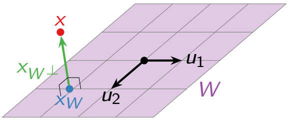
What is the projection of $x = \begin{pmatrix} -1.1 \\ 1.4 \\ 1.45 \end{pmatrix}$ onto $W = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1.1 \\ -.2 \end{pmatrix} \right\}$?

Answer: The basis is orthogonal, so

$$x_W = \mathrm{proj}_W \begin{pmatrix} -1.1 \\ 1.4 \\ 1.45 \end{pmatrix} = \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2} u_2$$

$$= \frac{(-1.1)(1)}{1^2} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \frac{(1.4)(1.1) + (1.45)(-.2)}{1.1^2 + (-.2)^2} \begin{pmatrix} 0 \\ 1.1 \\ -.2 \end{pmatrix}$$

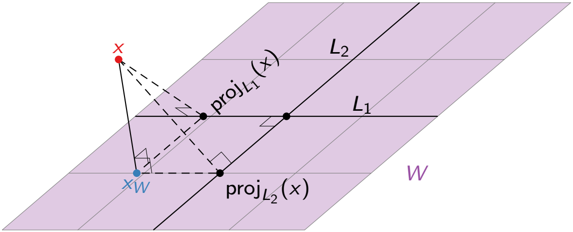
This turns out to be equal to $u_2 - 1.1u_1$.



Let W be a subspace of $\mathbf{R}^n$, and let $\{u_1, u_2, \ldots, u_m\}$ be an orthogonal basis for W. Let $L_i = \mathrm{Span}\{u_i\}$. Then

$$\mathrm{proj}_W(x) = \sum_{i=1}^m \frac{x \cdot u_i}{u_i \cdot u_i}\, u_i = \sum_{i=1}^m \mathrm{proj}_{L_i}(x).$$

So the orthogonal projection is formed by adding orthogonal projections onto perpendicular lines.



First we restate the property we've been using all along.

Best Approximation Theorem

Let W be a subspace of $\mathbf{R}^n$, and let x be a vector in $\mathbf{R}^n$. Then $y = \mathrm{proj}_W(x)$ is the closest point in W to x, in the sense that

$$\mathrm{dist}(x, y') \geq \mathrm{dist}(x, y) \quad \text{for all} \quad y' \text{ in } W.$$

We can think of orthogonal projection as a *transformation*:

$$\mathrm{proj}_W \colon \mathbf{R}^n \longrightarrow \mathbf{R}^n \qquad x \mapsto \mathrm{proj}_W(x).$$

Theorem

Let W be a subspace of $\mathbf{R}^n$.

1. proj_W is a *linear* transformation.
2. For every x in W, we have $\mathrm{proj}_W(x) = x$.
3. For every x in $W^\perp$, we have $\mathrm{proj}_W(x) = 0$.
4. The range of proj_W is W.

Poll

Let W be a subspace of $\mathbf{R}^n$.

Poll

Let A be the matrix for proj_W. What is/are the eigenvalue(s) of A?

A. 0 B. 1 C. -1 D. $0, 1$ E. $1, -1$ F. $0, -1$ G. $-1, 0, 1$

The 1-eigenspace is W.

The 0-eigenspace is $W^\perp$.

We have $\dim W + \dim W^\perp = n$, so that gives n linearly independent eigenvectors already.

So the answer is D.

What is the matrix for $\mathrm{proj}_W\colon \mathbf{R}^3 \to \mathbf{R}^3$, where

$$W = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}?$$

Answer: Recall how to compute the matrix for a linear transformation:

$$A = \begin{pmatrix} | & | & | \\ \mathrm{proj}_W(e_1) & \mathrm{proj}_W(e_2) & \mathrm{proj}_W(e_3) \\ | & | & | \end{pmatrix}.$$

We compute:

$$\mathrm{proj}_W(e_1) = \frac{e_1 \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{e_1 \cdot u_2}{u_2 \cdot u_2} u_2 = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5/6 \\ 1/3 \\ -1/6 \end{pmatrix}$$

$$\mathrm{proj}_W(e_2) = \frac{e_2 \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{e_2 \cdot u_2}{u_2 \cdot u_2} u_2 = 0 + \frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}$$

$$\mathrm{proj}_W(e_3) = \frac{e_3 \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{e_3 \cdot u_2}{u_2 \cdot u_2} u_2 = -\frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/6 \\ 1/3 \\ 5/6 \end{pmatrix}$$

Therefore $A = \begin{pmatrix} 5/6 & 1/3 & -1/6 \\ 1/3 & 1/3 & 1/3 \\ -1/6 & 1/3 & 5/6 \end{pmatrix}$.

Let W be an m-dimensional subspace of $\mathbf{R}^n$, let $\text{proj}_W : \mathbf{R}^n \to W$ be the projection, and let A be the matrix for proj_L.

Fact 1: A is diagonalizable with eigenvalues 0 and 1; it is similar to the diagonal matrix with m ones and $n - m$ zeros on the diagonal.

Why? Let $v_1, v_2, \ldots, v_m$ be a basis for W, and let $v_{m+1}, v_{m+2}, \ldots, v_n$ be a basis for $W^\perp$. These are (linearly independent) eigenvectors with eigenvalues 1 and 0, respectively, and they form a basis for $\mathbf{R}^n$ because there are n of them.

Example: If W is a plane in $\mathbf{R}^3$, then A is similar to projection onto the xy-plane:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Fact 2: $A^2 = A$.

Why? Projecting twice is the same as projecting once:

$$\text{proj}_W \circ \text{proj}_W = \text{proj}_W \implies A \cdot A = A.$$

What is the distance from e_1 to $W = \text{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$?

Answer: The closest point on W to e_1 is $\text{proj}_W(e_1) = \begin{pmatrix} 5/6 \\ 1/3 \\ -1/6 \end{pmatrix}$.

The distance from e_1 to this point is

$$\begin{aligned}
\text{dist}\big(e_1, \text{proj}_W(e_1)\big) &= \|(e_1)_{W^\perp}\| \\
&= \left\| \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 5/6 \\ 1/3 \\ -1/6 \end{pmatrix} \right\| \\
&= \left\| \begin{pmatrix} 1/6 \\ -1/3 \\ -1/6 \end{pmatrix} \right\| \\
&= \sqrt{(1/6)^2 + (-1/3)^2 + (-1/6)^2} \\
&= \frac{1}{\sqrt{6}}.
\end{aligned}$$