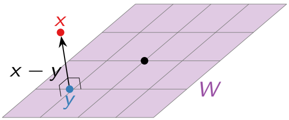
Section 6.2

Orthogonal Sets

Best Approximation

Due to measurement error, the measured x is not actually in the subspace it must lie on (*for theoretical reasons*).



Best approximation: y is the *closest point* to x on W.

Replace x with its orthogonal projection y onto W.

How do you know that y is the closest point?

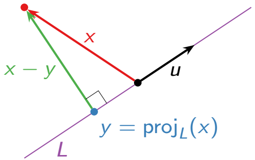
Orthogonal Projection onto a Line

Theorem

Let $L = \text{Span}\{u\}$ *be a line* in $\mathbf{R}^n$, and let x be in $\mathbf{R}^n$. The closest point to x on L is the point

$$\text{proj}_L(x) = \frac{x \cdot u}{u \cdot u}\, u.$$

This point is called the **orthogonal projection of x onto L.**



Choose term 'ortogonal' because $x - y$ *is in* $L^\perp$. That is, $u \cdot (x - y) = 0$:

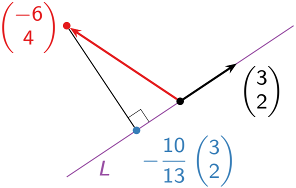
$$u \cdot (x - y) = u \cdot \left(x - \frac{x \cdot u}{u \cdot u} u\right) = u \cdot x - \frac{x \cdot u}{u \cdot u}(u \cdot u) = u \cdot x - x \cdot u = 0.$$

Orthogonal Projection onto a Line

Example

Compute the *orthogonal projection* of $x = \begin{pmatrix} -6 \\ 4 \end{pmatrix}$ onto the line *L spanned by* $u = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$.

$$y = \mathrm{proj}_L(x) = \frac{x \cdot u}{u \cdot u} u = \frac{-18 + 8}{9 + 4} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = -\frac{10}{13} \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$



Orthogonal Sets

Definition

A set of *nonzero* vectors is **orthogonal** if each pair of vectors is orthogonal. Such set is **orthonormal** if, in addition, each vector is a *unit vector*.

Example: $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\}$ is an orthogonal set.

Check:

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = 0 \qquad \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = 0 \qquad \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = 0.$$

Orthogonal bases

Linearly independent

An *orthogonal set of vectors* $\mathcal{B} = \{u_1, u_2, \ldots, u_m\}$ is linearly independent. Therefore $\mathcal{B}$ **forms a basis** for $W = \operatorname{Span}\mathcal{B}$.

Why?

Suppose $\{u_1, u_2, \ldots, u_m\}$ *is orthogonal* and that

$$c_1 u_1 + c_2 u_2 + \cdots + c_m u_m = 0$$

Now **dot-multiply by u_1**. We see that c_1 *must be zero*:

$$0 = u_1 \cdot (c_1 u_1 + c_2 u_2 + \cdots + c_m u_m) = c_1(u_1 \cdot u_1) + 0 + 0 + \cdots + 0.$$

Similarly for the other c_i's (there is only trivial solution). Hence the set is linearly independent.

$\mathcal{B}$-coordinates for Orthogonal bases

Theorem

Let $\mathcal{B} = \{u_1, u_2, \ldots, u_m\}$ be an orthogonal set, and let x be a vector in $W = \operatorname{Span}\mathcal{B}$. Then

$$x = \sum_{i=1}^{m} \frac{x \cdot u_i}{u_i \cdot u_i}\, u_i = \frac{x \cdot u_1}{u_1 \cdot u_1}\, u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2}\, u_2 + \cdots + \frac{x \cdot u_m}{u_m \cdot u_m}\, u_m.$$

An advantage

For orthogonal bases, is it's *easy to compute the $\mathcal{B}$-coordinates* of a vector x in W:

$$\left(\frac{x \cdot u_1}{u_1 \cdot u_1}, \frac{x \cdot u_2}{u_2 \cdot u_2}, \ldots, \frac{x \cdot u_m}{u_m \cdot u_m}\right).$$

Why? If $x = c_1 u_1 + c_2 u_2 + \cdots + c_m u_m$, then

$$x \cdot u_1 = c_1(u_1 \cdot u_1) + 0 + \cdots + 0 \implies c_1 = \frac{x \cdot u_1}{u_1 \cdot u_1}.$$

Similarly for the other c_i's.

Orthogonal Bases

Geometric reason

Theorem

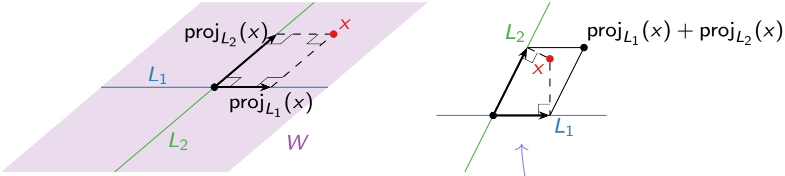
Let $\mathcal{B} = \{u_1, u_2, \ldots, u_m\}$ be an orthogonal set, and let x be a vector in $W = \operatorname{Span} \mathcal{B}$. Then

$\operatorname{proj}_{L_2}(u_2)$

$$x = \sum_{i=1}^{m} \frac{x \cdot u_i}{u_i \cdot u_i}\, u_i = \frac{x \cdot u_1}{u_1 \cdot u_1}\, u_1 + \boxed{\frac{x \cdot u_2}{u_2 \cdot u_2}\, u_2} + \cdots + \frac{x \cdot u_m}{u_m \cdot u_m}\, u_m.$$

If L_i is the line spanned by u_i, then this says

$$x = \operatorname{proj}_{L_1}(x) + \operatorname{proj}_{L_2}(x) + \cdots + \operatorname{proj}_{L_m}(x).$$



Warning: This only works for an *orthogonal* basis.

Problem: Find the $\mathcal{B}$-coordinates of $x = \binom{0}{3}$, where

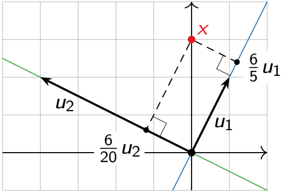
$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -4 \\ 2 \end{pmatrix} \right\}.$$

Old way:

$$\left(\begin{array}{cc|c} 1 & -4 & 0 \\ 2 & 2 & 3 \end{array}\right) \xrightarrow{\text{rref}} \left(\begin{array}{cc|c} 1 & 0 & 6/5 \\ 0 & 1 & 6/20 \end{array}\right) \implies [x]_{\mathcal{B}} = \begin{pmatrix} 6/5 \\ 6/20 \end{pmatrix}.$$

New way: Exploit that $\mathcal{B}$ is an *orthogonal basis*.

$$x = \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2} u_2 = \frac{3 \cdot 2}{1^2 + 2^2} u_1 + \frac{3 \cdot 2}{(-4)^2 + 2^2} u_2 = \frac{6}{5} u_1 + \frac{6}{20} u_2.$$



$$\implies [x]_{\mathcal{B}} = \begin{pmatrix} 6/5 \\ 6/20 \end{pmatrix}.$$

Orthogonal Bases

Example

Problem: Find the *$\mathcal{B}$-coordinates* of $x = (6, 1, -8)$ where

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\}.$$

Answer: Check that $\mathcal{B}$ is *orthogonal basis*, then

$$\begin{aligned} [x]_{\mathcal{B}} &= \left(\frac{x \cdot u_1}{u_1 \cdot u_1}, \frac{x \cdot u_2}{u_2 \cdot u_2}, \frac{x \cdot u_3}{u_3 \cdot u_3} \right) \\ &= \left(\frac{6 \cdot 1 + 1 \cdot 1 - 8 \cdot 1}{1^2 + 1^2 + 1^2}, \frac{6 \cdot 1 + 1 \cdot (-2) - 8 \cdot 1}{1^2 + (-2)^2 + 1^2}, \frac{6 \cdot 1 + 1 \cdot 0 + (-8) \cdot (-1)}{1^2 + 0^2 + (-1)^2} \right) \\ &= \left(-\frac{1}{3}, -\frac{2}{3}, 7 \right). \end{aligned}$$

Check:

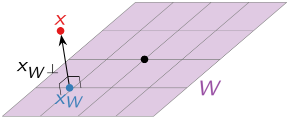
$$\begin{pmatrix} 6 \\ 1 \\ -8 \end{pmatrix} = -\frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{3} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + 7 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

✔

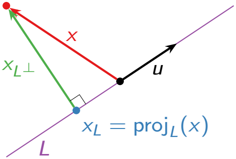
Section 6.3

Orthogonal Projections

Motivation



Example with a line: The closest point to x in L is $\text{proj}_L(x) = \frac{x \cdot u}{u \cdot u} u$



Let $u = \binom{3}{2}$ and let $L = \text{Span}\{u\}$. Let $x = \binom{-6}{4}$. In this case,

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

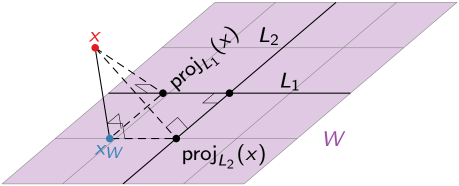
Definition

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

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

Note: If $L_i = \text{Span}\{u_i\}$. Then $\frac{x \cdot u_i}{u_i \cdot u_i}\, u_i = \text{proj}_{L_i}(x)$.

The orthogonal projection is formed by adding orthogonal projections onto perpendicular lines.



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.

The following is *the property we wanted* 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

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

Best approximation

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 $W^\perp$.

Theorem

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

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

Why? Let $y = \text{proj}_W(x)$. 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:

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

Poll

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

Poll

Let A be the *matrix for* proj_W. What are *all the eigenvalues* 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; and 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 A **be the matrix for** proj_W, where W is an m-dimensional subspace of $\mathbf{R}^n$.

Facts:

1. A *is diagonalizable* with eigenvalues 0 and 1;
2. it is similar to the *diagonal matrix with m ones and $n - m$ zeros* on the diagonal, and
3. $A^2 = A$.

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

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

Why 1-2? 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.

Why 3? Projecting twice is the same as projecting once:

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

Orthogonal Projections

Minimum distance

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

Answer: From e_1 to its closest point on W:

$$\text{dist}(e_1, \text{proj}_W(e_1)) = \|(e_1)_{W^\perp}\|.$$

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