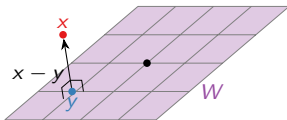


# Section 6.4

## The Gram–Schmidt Process

## Motivation: Best Approximation

Suppose you measure a data point  $x$  which you know for theoretical reasons must lie on a subspace  $W$ .



Due to measurement error, though, the measured  $x$  is not actually in  $W$ . Best approximation:  $y$  is the *closest* point to  $x$  on  $W$ .

How do you know that  $y$  is the closest point? The vector from  $y$  to  $x$  is orthogonal to  $W$ : it is in the *orthogonal complement*  $W^\perp$ .

Note  $x = y + (x - y)$ , where  $y$  is in  $W$  and  $x - y$  is in  $W^\perp$ . Last time we called this the *orthogonal decomposition* of  $x$ :

$$x = x_W + x_{W^\perp} \quad x_W = y \quad x_{W^\perp} = x - y.$$

# Orthogonal Decomposition

## Review

**Recall:** If  $W$  is a subspace of  $\mathbf{R}^n$ , its **orthogonal complement** is

$$W^\perp = \{v \text{ in } \mathbf{R}^n \mid v \text{ is perpendicular to every vector in } W\}$$

## Theorem

Every vector  $x$  in  $\mathbf{R}^n$  can be written as

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

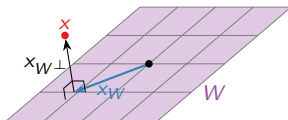
for unique vectors  $x_W$  in  $W$  and  $x_{W^\perp}$  in  $W^\perp$ .

The equation  $x = x_W + x_{W^\perp}$  is called the **orthogonal decomposition** of  $x$  (with respect to  $W$ ).

The vector  $x_W$  is the closest vector to  $x$  on  $W$ .

[interactive 1]

[interactive 2]



# Orthogonal Projections

## Review

How do you compute  $x_W$ ? (Note  $x_{W^\perp} = x - x_W$ .)

**Recall:** a set of nonzero vectors  $\{u_1, u_2, \dots, u_m\}$  is **orthogonal** if  $u_i \cdot u_j = 0$  when  $i \neq j$ : each vector is perpendicular to the others.

### Definition

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

$$\text{proj}_W(x) \stackrel{\text{def}}{=} \sum_{i=1}^m \frac{x \cdot u_i}{u_i \cdot u_i} u_i = \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2} u_2 + \dots + \frac{x \cdot u_n}{u_n \cdot u_n} u_n.$$

[interactive]

### Theorem

Let  $W$  be a subspace of  $\mathbf{R}^n$ , and let  $x$  be a vector in  $\mathbf{R}^n$ . Then

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

In particular,  $\text{proj}_W(x)$  is the closest point to  $x$  in  $W$ .

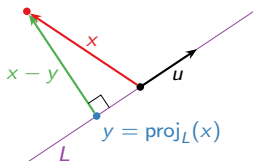
# Orthogonal Projection onto a Line

## Review

The formula for orthogonal projections is simple when  $W$  is a *line*.

Let  $L = \text{Span}\{u\}$  be a line in  $\mathbf{R}^n$ , and let  $x$  be in  $\mathbf{R}^n$ . The orthogonal projection of  $x$  onto  $L$  is the point

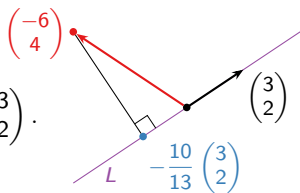
$$\text{proj}_L(x) = \frac{x \cdot u}{u \cdot u} u.$$



[interactive]

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

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



# Orthogonal Projections

## Properties

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

$$\text{proj}_W: \mathbf{R}^n \longrightarrow \mathbf{R}^n \quad x \mapsto \text{proj}_W(x).$$

### Theorem

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

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

Let  $W$  be a subspace with orthogonal basis  $\mathcal{B} = \{u_1, u_2, \dots, u_m\}$ .

For  $x$  in  $W$  we have  $\text{proj}_W(x) = x$ , so

$$\begin{aligned} x = \text{proj}_W(x) &= \sum_{i=1}^m \frac{x \cdot u_i}{u_i \cdot u_i} u_i = \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2} u_2 + \dots + \frac{x \cdot u_n}{u_n \cdot u_n} u_n \\ \implies [x]_{\mathcal{B}} &= \left( \frac{x \cdot u_1}{u_1 \cdot u_1}, \frac{x \cdot u_2}{u_2 \cdot u_2}, \dots, \frac{x \cdot u_m}{u_m \cdot u_m} \right). \quad \text{[interactive]} \end{aligned}$$

## A Non-Orthogonal Basis

**Important:** Orthogonal projections require an *orthogonal* basis!

**Non-Example:** Consider the basis  $\mathcal{B} = \{v_1, v_2\}$  of  $\mathbf{R}^2$ , where

$$v_1 = \begin{pmatrix} 2 \\ -1/2 \end{pmatrix} \quad v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

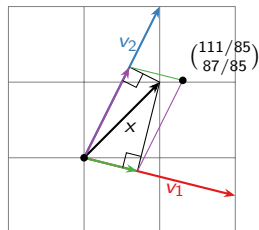
This is not orthogonal:  $\begin{pmatrix} 2 \\ -1/2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 1 \neq 0$ .

Let's try to compute  $x = \text{proj}_{\mathbf{R}^2}(x)$  for  $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$  using the basis  $\{v_1, v_2\}$ :

$$x = \text{proj}_{\mathbf{R}^2}(x) = \frac{x \cdot v_1}{v_1 \cdot v_1} v_1 + \frac{x \cdot v_2}{v_2 \cdot v_2} v_2 = \frac{3/2}{17/4} \begin{pmatrix} 2 \\ -1/2 \end{pmatrix} + \frac{3}{5} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 111/85 \\ 87/85 \end{pmatrix} \quad \times$$

This does not work!

[interactive]



All of the procedures we learned in §§6.2–6.3 require an *orthogonal* basis  $\{u_1, u_2, \dots, u_m\}$ .

- ▶ Finding the orthogonal projection of a vector  $x$  onto the span  $W$  of  $u_1, u_2, \dots, u_m$ :

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

- ▶ Finding the orthogonal decomposition of  $x$ :

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

- ▶ Finding the  $\mathcal{B}$ -coordinates of  $x$ :

$$[x]_{\mathcal{B}} = \left( \frac{x \cdot u_1}{u_1 \cdot u_1}, \frac{x \cdot u_2}{u_2 \cdot u_2}, \dots, \frac{x \cdot u_m}{u_m \cdot u_m} \right).$$

**Problem:** What if your basis isn't orthogonal?

**Solution:** The Gram–Schmidt process: take any basis and make it orthogonal.



# The Gram–Schmidt Process

## Procedure

### The Gram–Schmidt Process

Let  $\{v_1, v_2, \dots, v_m\}$  be a basis for a subspace  $W$  of  $\mathbf{R}^n$ . Define:

$$1. \quad u_1 = v_1$$

$$2. \quad u_2 = v_2 - \text{proj}_{\text{Span}\{u_1\}}(v_2) = v_2 - \frac{v_2 \cdot u_1}{u_1 \cdot u_1} u_1$$

$$3. \quad u_3 = v_3 - \text{proj}_{\text{Span}\{u_1, u_2\}}(v_3) = v_3 - \frac{v_3 \cdot u_1}{u_1 \cdot u_1} u_1 - \frac{v_3 \cdot u_2}{u_2 \cdot u_2} u_2$$

$\vdots$

$$m. \quad u_m = v_m - \text{proj}_{\text{Span}\{u_1, u_2, \dots, u_{m-1}\}}(v_m) = v_m - \sum_{i=1}^{m-1} \frac{v_m \cdot u_i}{u_i \cdot u_i} u_i$$

Then  $\{u_1, u_2, \dots, u_m\}$  is an *orthogonal* basis for the same subspace  $W$ .

### Remark

In fact, for every  $i$  between 1 and  $n$ , the set  $\{u_1, u_2, \dots, u_i\}$  is an orthogonal basis for  $\text{Span}\{v_1, v_2, \dots, v_i\}$ .

# The Gram–Schmidt Process

## Two vectors

Find an orthogonal basis  $\{u_1, u_2\}$  for  $W = \text{Span}\{v_1, v_2\}$ , where

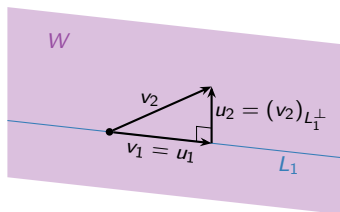
$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Run Gram–Schmidt:

$$1. \quad u_1 = v_1 \quad 2. \quad u_2 = v_2 - \frac{v_2 \cdot u_1}{u_1 \cdot u_1} u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Why does this work?

- ▶ First we take  $u_1 = v_1$ .
- ▶ Now we're sad because  $u_1 \cdot v_2 \neq 0$ , so we can't take  $u_2 = v_2$ .
- ▶ Fix: let  $L_1 = \text{Span}\{u_1\}$ , and let  $u_2 = (v_2)_{L_1^\perp} = v_2 - \text{proj}_{L_1}(v_2)$ .
- ▶ By construction,  $u_1 \cdot u_2 = 0$ , because  $L_1 \perp u_2$ .



**Important:**  $\text{Span}\{u_1, u_2\} = \text{Span}\{v_1, v_2\} = W$ : this is an *orthogonal* basis for the *same* subspace.

# The Gram–Schmidt Process

## Three vectors

Find an orthogonal basis  $\{u_1, u_2, u_3\}$  for  $W = \text{Span}\{v_1, v_2, v_3\} = \mathbf{R}^3$ , where

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad v_3 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}.$$

Run Gram–Schmidt:

1.  $u_1 = v_1$

2.  $u_2 = v_2 - \frac{v_2 \cdot u_1}{u_1 \cdot u_1} u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

3.  $u_3 = v_3 - \frac{v_3 \cdot u_1}{u_1 \cdot u_1} u_1 - \frac{v_3 \cdot u_2}{u_2 \cdot u_2} u_2$   
 $= \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} - \frac{4}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{1} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$

**Important:**  $\text{Span}\{u_1, u_2, u_3\} = \text{Span}\{v_1, v_2, v_3\} = W$ : this is an *orthogonal* basis for the *same* subspace.

# The Gram–Schmidt Process

Three vectors, continued

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \xrightarrow{\text{G-S}} u_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, u_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

Why does this work?

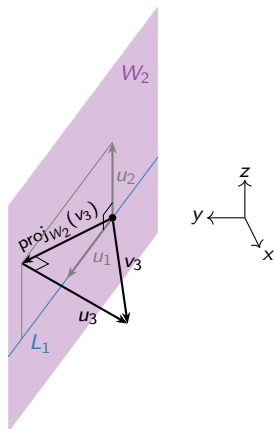
- ▶ Once we have  $u_1$  and  $u_2$ , then we're sad because  $v_3$  is not orthogonal to  $u_1$  and  $u_2$ .
- ▶ Fix: let  $W_2 = \text{Span}\{u_1, u_2\}$ , and let  $u_3 = (v_3)_{W_2^\perp} = v_3 - \text{proj}_{W_2}(v_3)$ .
- ▶ By construction,  $u_1 \cdot u_3 = 0 = u_2 \cdot u_3$  because  $W_2 \perp u_3$ .

Check:

$$u_1 \cdot u_2 = 0 \quad \checkmark \checkmark$$

$$u_1 \cdot u_3 = 0 \quad \checkmark \checkmark$$

$$u_2 \cdot u_3 = 0 \quad \checkmark \checkmark$$



# The Gram–Schmidt Process

Three vectors in  $\mathbb{R}^4$

Find an orthogonal basis  $\{u_1, u_2, u_3\}$  for  $W = \text{Span}\{v_1, v_2, v_3\}$ , where

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad v_2 = \begin{pmatrix} -1 \\ 4 \\ 4 \\ -1 \end{pmatrix} \quad v_3 = \begin{pmatrix} 4 \\ -2 \\ -2 \\ 0 \end{pmatrix}.$$

Run Gram–Schmidt:

1.  $u_1 = v_1$

2.  $u_2 = v_2 - \frac{v_2 \cdot u_1}{u_1 \cdot u_1} u_1 = \begin{pmatrix} -1 \\ 4 \\ 4 \\ -1 \end{pmatrix} - \frac{6}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -5/2 \\ 5/2 \\ 5/2 \\ -5/2 \end{pmatrix}$

3.  $u_3 = v_3 - \frac{v_3 \cdot u_1}{u_1 \cdot u_1} u_1 - \frac{v_3 \cdot u_2}{u_2 \cdot u_2} u_2$   
 $= \begin{pmatrix} 4 \\ -2 \\ -2 \\ 0 \end{pmatrix} - \frac{0}{24} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{-20}{25} \begin{pmatrix} -5/2 \\ 5/2 \\ 5/2 \\ -5/2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \\ -2 \end{pmatrix}$

Poll (we didn't get to do it in class, so it will not be counted)

What happens if you try to run Gram–Schmidt on a linearly dependent set of vectors  $\{v_1, v_2, \dots, v_m\}$ ?

- A. You get an inconsistent equation.
- B. For some  $i$  you get  $u_i = u_{i-1}$ .
- C. For some  $i$  you get  $u_i = 0$ .
- D. You create a rift in the space-time continuum.

If  $\{v_1, v_2, \dots, v_m\}$  is linearly dependent, then some  $v_i$  is in  $\text{Span}\{v_1, v_2, \dots, v_{i-1}\} = \text{Span}\{u_1, u_2, \dots, u_{i-1}\}$ .

This means

$$v_i = \text{proj}_{\text{Span}\{u_1, u_2, \dots, u_{i-1}\}}(v_i)$$

$$\implies u_i = v_i - \text{proj}_{\text{Span}\{u_1, u_2, \dots, u_{i-1}\}}(v_i) = 0.$$

In this case, you can simply discard  $u_i$  and  $v_i$  and continue: so Gram–Schmidt produces an orthogonal basis from any spanning set!

## Summary

- ▶ We like orthogonal bases because they let us compute orthogonal projections.
- ▶ The Gram–Schmidt process turns an arbitrary basis into an orthogonal basis.