Announcements
Monday, November 27

- WeBWorK 6.1, 6.2, 6.3 are due Wednesday at 11:59pm.

- WeBWorK 6.4, 6.5 are posted and will be covered on the final, but they are not graded.

- No quiz on Friday! But this is the only recitation on chapter 6.

- My office is Skiles 244. Rabin office hours are Monday, 1–3pm and Tuesday, 9–11am.
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 $x$:

$$x = x_W + x_{W^\perp} \quad x_W = y \quad x_{W^\perp} = x - y.$$
Recall: If $W$ is a subspace of $\mathbb{R}^n$, its **orthogonal complement** is

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

**Theorem**

Every vector $x$ in $\mathbb{R}^n$ 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 $\{u_1, u_2, \ldots, u_m\}$ 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 $\mathbb{R}^n$, 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

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

Let $x$ be a vector and let $x = x_W + x_{W\perp}$ be its orthogonal decomposition with respect to a subspace $W$. The following vectors are the same:

- $x_W$
- $\text{proj}_W(x)$
- The closest vector to $x$ on $W$
Orthogonal Projection onto a Line

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

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

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

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

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

**Theorem**

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

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 $B = \{u_1, u_2, \ldots, u_m\}$.

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

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

$$\implies [x]_B = \left( \frac{x \cdot u_1}{u_1 \cdot u_1}, \frac{x \cdot u_2}{u_2 \cdot u_2}, \ldots, \frac{x \cdot u_m}{u_m \cdot u_m} \right).$$
Important: Orthogonal projections require an *orthogonal* basis!

Non-Example: Consider the basis $B = \{v_1, v_2\}$ of $\mathbb{R}^2$, 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 $x = \text{proj}_{\mathbb{R}^2}(x)$ for $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ using the basis $\{v_1, v_2\}$:

$$x = \text{proj}_{\mathbb{R}^2}(x) =$$

This does not work!

[interactive]  (compare [orthogonal basis])
All of the procedures we learned in §§6.2–6.3 require an *orthogonal* basis $\{u_1, u_2, \ldots, u_m\}$.

- Finding the orthogonal projection of a vector $x$ onto the span $W$ of $u_1, u_2, \ldots, u_m$:
  $$\text{proj}_W(x) = \sum_{i=1}^{m} \frac{x \cdot u_i}{u_i \cdot u_i} u_i.$$  

- Finding the orthogonal decomposition of $x$:
  $$x = \text{proj}_W(x) + x_{W\perp}.$$  

- Finding the $B$-coordinates of $x$:
  $$[x]_B = \left( \frac{x \cdot u_1}{u_1 \cdot u_1}, \frac{x \cdot u_2}{u_2 \cdot u_2}, \ldots, \frac{x \cdot u_m}{u_m \cdot u_m} \right).$$

**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**

Let \( \{v_1, v_2, \ldots, v_m\} \) be a basis for a subspace \( W \) of \( \mathbb{R}^n \). Define:

1. \( u_1 = v_1 \)
2. \( u_2 = v_2 - \text{proj}_{\text{Span}\{u_1\}}(v_2) = v_2 - \frac{v_2 \cdot u_1}{u_1 \cdot u_1} u_1 \)
3. \( u_3 = v_3 - \text{proj}_{\text{Span}\{u_1,u_2\}}(v_3) = v_3 - \frac{v_3 \cdot u_1}{u_1 \cdot u_1} u_1 - \frac{v_3 \cdot u_2}{u_2 \cdot u_2} u_2 \)

\[ \vdots \]

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

Then \( \{u_1, u_2, \ldots, u_m\} \) is an **orthogonal** basis for the same subspace \( W \).

**Remark**

In fact, for every \( i \) between 1 and \( n \), the set \( \{u_1, u_2, \ldots, u_i\} \) is an orthogonal basis for \( \text{Span}\{v_1, v_2, \ldots, v_i\} \).
The Gram–Schmidt Process

Two vectors

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

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

**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

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

\[
\begin{align*}
v_1 &= \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, & v_2 &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, & v_3 &= \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}.
\end{align*}
\]

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

\[ \mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \]

\( \text{G–S} \) \[ \mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \]

Why does this work?

\[ \text{Once we have } \mathbf{u}_1 \text{ and } \mathbf{u}_2, \text{ then we're sad because } \mathbf{v}_3 \text{ is not orthogonal to } \mathbf{u}_1 \text{ and } \mathbf{u}_2. \]

\[ \text{Fix: let } W_2 = \text{Span}\{ \mathbf{u}_1, \mathbf{u}_2 \}, \text{ and let } \mathbf{u}_3 = (\mathbf{v}_3)_{W_2} = \mathbf{v}_3 - \text{proj}_{W_2}(\mathbf{v}_3). \]

\[ \text{By construction, } \mathbf{u}_1 \cdot \mathbf{u}_3 = 0 = \mathbf{u}_2 \cdot \mathbf{u}_3 \text{ because } W_2 \perp \mathbf{u}_3. \]

Check:

\[ \mathbf{u}_1 \cdot \mathbf{u}_2 = 0 \]

\[ \mathbf{u}_1 \cdot \mathbf{u}_3 = 0 \]

\[ \mathbf{u}_2 \cdot \mathbf{u}_3 = 0 \]
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

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

What happens if you try to run Gram–Schmidt on a linearly independent set of vectors \{v_1, v_2, \ldots, 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.

Poll

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

This means \(v_i = \text{proj}_{\text{Span}\{u_1, u_2, \ldots, u_{i-1}\}}(v_i) \Rightarrow u_i = v_i - \text{proj}_{\text{Span}\{u_1, u_2, \ldots, u_{i-1}\}}(v_i) = 0\).

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!
We like orthogonal bases because they let us compute orthogonal projections.

The Gram–Schmidt process turns an arbitrary basis into an orthogonal basis.