The final exam will take place this Friday, 12/9, 8am–10:50am in this room.

Please fill out the CIOS form online. (Current response rate: 76%)
- It is important for me to get responses from most of the class: I use these for preparing future iterations of this course.
- If we get an 80% response rate before the final, I’ll drop the two lowest quiz grades instead of one.

WeBWorK assignments 6.4 and 6.5 are posted, and the solutions are visible. (For practice only; not scored.)

Extra office hours: today 1–3pm, tomorrow 2–4pm, Thursday 11am–4pm, and by appointment, in Skiles 221.
- As always, TAs’ office hours are posted on the website.
- Math Lab is also a good place to visit.
Review for Final Exam

Selected Topics
Orthogonal Sets

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

Example: \( \mathcal{B}_1 = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \) is not orthogonal.

Example: \( \mathcal{B}_2 = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \) is orthogonal but not orthonormal.

Example: \( \mathcal{B}_3 = \left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} \) is orthonormal.

To go from an orthogonal set \( \{u_1, u_2, \ldots, u_m\} \) to an orthonormal set, replace each \( u_i \) with \( u_i/\|u_i\| \).

Theorem
An orthogonal set is linearly independent. In particular, it is a basis for its span.
Let $W$ be a subspace of $\mathbb{R}^n$, and let $B = \{u_1, u_2, \ldots, u_m\}$ be an orthogonal basis for $W$. The **orthogonal projection** of a vector $x$ onto $W$ is

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

This is the closest vector to $x$ that lies on $W$. In other words, the difference $x - \text{proj}_W(x)$ is perpendicular to $W$: it is in $W^\perp$. Notation:

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

So $x_W$ is in $W$, $x_{W^\perp}$ is in $W^\perp$, and $x = x_W + x_{W^\perp}$. 
Orthogonal Projection

Special cases

**Special case:** If \( x \) is in \( W \), then \( x = \text{proj}_W(x) \), so

\[
x = \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2} u_2 + \cdots + \frac{x \cdot u_m}{u_m \cdot u_m} u_m.
\]

In other words, the \( B \)-coordinates of \( x \) are

\[
\left( \frac{x \cdot u_1}{u_1 \cdot u_1}, \frac{x \cdot u_2}{u_2 \cdot u_2}, \ldots, \frac{x \cdot u_m}{u_m \cdot u_m} \right),
\]

where \( B = \{u_1, u_2, \ldots, u_m\} \), a basis for \( W \).

**Special case:** If \( W = L \) is a line, then \( L = \text{Span}\{u\} \) for some nonzero vector \( u \), and

\[
\text{proj}_L(x) = \frac{x \cdot u}{u \cdot u} u
\]
Let $W$ be a subspace of $\mathbb{R}^n$.

**Theorem**
The orthogonal projection $\text{proj}_W$ is a *linear* transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$. Its range is $W$.

If $A$ is the matrix for $\text{proj}_W$, then $A^2 = A$ because projecting twice is the same as projecting once: $\text{proj}_W \circ \text{proj}_W = \text{proj}_W$.

**Theorem**
The only eigenvalues of $A$ are 1 and 0.

**Why?**

$$A\mathbf{v} = \lambda \mathbf{v} \implies A^2 \mathbf{v} = A(A\mathbf{v}) = A(\lambda \mathbf{v}) = \lambda (A\mathbf{v}) = \lambda^2 \mathbf{v}.$$  

So if $\lambda$ is an eigenvalue of $A$, then $\lambda^2$ is an eigenvalue of $A^2$. But $A^2 = A$, so $\lambda^2 = \lambda$, and hence $\lambda = 0$ or 1.

The 1-eigenspace of $A$ is $W$, and the 0-eigenspace is $W^\perp$. 
The Gram–Schmidt Process

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

... 

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

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

In fact, for each \( i \),

\[ \text{Span}\{u_1, u_2, \ldots, u_i\} = \text{Span}\{v_1, v_2, \ldots, v_i\}. \]

Note if \( v_i \) is in \( \text{Span}\{v_1, v_2, \ldots, v_{i-1}\} = \text{Span}\{u_1, u_2, \ldots, u_{i-1}\} \), then

\( v_i = \text{proj}_{\text{Span}\{u_1, u_2, \ldots, u_{i-1}\}}(v_i) \), so \( u_i = 0 \). So this also detects linear dependence.
QR Factorization

Let $A$ be a matrix with linearly independent columns. Then

$$A = QR$$

where $Q$ has orthonormal columns and $R$ is upper-triangular with positive diagonal entries.

Step 1: Let $v_1, v_2, \ldots, v_m$ be the columns of $A$. Run Gram–Schmidt on \{\(v_1, v_2, \ldots, v_m\)\} to get an orthogonal basis \{\(u_1, u_2, \ldots, u_m\)\}, and solve for each $v_i$ in terms of $u_1, u_2, \ldots, u_i$.

Step 2: Put the resulting equations in matrix form to get $A = \hat{Q}\hat{R}$ where

$$A = \begin{pmatrix} v_1 & v_2 & \cdots & v_m \end{pmatrix} \quad \hat{Q} = \begin{pmatrix} u_1 & u_2 & \cdots & u_m \end{pmatrix}$$

and $\hat{R}$ contains the coefficients from $v_i = \text{(linear combination of } u_1, u_2, \ldots, u_{i-1})$ in the columns.

Step 3: Scale each column of $\hat{Q}$ by its length to get a matrix with orthonormal columns, and scale each row of $\hat{R}$ by the opposite factor to get $Q$ and $R$, respectively.
Find the QR factorization of $A = \begin{pmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & -2 \\ 1 & -1 & 0 \end{pmatrix}$.

**Step 1:** Let $v_1, v_2, v_3$ be the columns. Run Gram–Schmidt and solve for $v_1, v_2, v_3$ in terms of $u_1, u_2, u_3$:

$$u_1 = v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad v_1 = u_1$$

$$u_2 = v_2 - \frac{v_2 \cdot u_1}{u_1 \cdot u_1} u_1 = v_2 - \frac{3}{2} u_1 = \begin{pmatrix} -5/2 \\ 5/2 \\ 5/2 \\ -5/2 \end{pmatrix} \quad v_2 = \frac{3}{2} u_1 + u_2$$

$$u_3 = v_3 - \frac{v_3 \cdot u_1}{u_1 \cdot u_1} u_1 - \frac{v_3 \cdot u_2}{u_2 \cdot u_2} u_2 = v_3 + \frac{4}{5} u_2 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ -2 \end{pmatrix} \quad v_3 = -\frac{4}{5} u_2 + u_3$$
$v_1 = 1 \ u_1 \quad v_2 = \frac{3}{2} \ u_1 + 1 \ u_2 \quad v_3 = 0 \ u_1 - \frac{4}{5} \ u_2 + 1 \ u_3$

Step 2: write $A = \hat{Q} \hat{R}$, where $\hat{Q}$ has orthogonality columns $u_1, u_2, u_3$ and $\hat{R}$ is upper-triangular with 1s on the diagonal.

$\hat{Q} = \begin{pmatrix}
1 & u_1 \\
& u_2 \\
& & u_3
\end{pmatrix} = \begin{pmatrix}
1 & -\frac{5}{2} & 2 \\
1 & \frac{5}{2} & 0 \\
1 & -\frac{5}{2} & -2
\end{pmatrix}$

$\hat{R} = \begin{pmatrix}
1 & \frac{3}{2} & 0 \\
0 & 1 & -\frac{4}{5} \\
0 & 0 & 1
\end{pmatrix}$
**QR Factorization**

Example, continued

\[
A = \hat{Q}\hat{R} \quad \hat{Q} = \begin{pmatrix}
1 & -5/2 & 2 \\
1 & 5/2 & 0 \\
1 & 5/2 & 0 \\
1 & -5/2 & -2
\end{pmatrix} \quad \hat{R} = \begin{pmatrix}
1 & 3/2 & 0 \\
0 & 1 & -4/5 \\
0 & 0 & 1
\end{pmatrix}
\]

**Step 3:** normalize the columns of \( \hat{Q} \) and the rows of \( \hat{R} \) to get \( Q \) and \( R \):

\[
Q = \begin{pmatrix}
\frac{1}{\|u_1\|} & \frac{1}{\|u_2\|} & \frac{1}{\|u_3\|}
\end{pmatrix} = \begin{pmatrix}
1/2 & -1/2 & 1/\sqrt{2} \\
1/2 & 1/2 & 0 \\
1/2 & 1/2 & 0 \\
1/2 & -1/2 & -1/\sqrt{2}
\end{pmatrix}
\]

\[
R = \begin{pmatrix}
1 \cdot \|u_1\| & 3/2 \cdot \|u_1\| & 0 \cdot \|u_1\| \\
0 & 1 \cdot \|u_2\| & -4/5 \cdot \|u_2\| \\
0 & 0 & 1 \cdot \|u_3\|
\end{pmatrix} = \begin{pmatrix}
2 & 3 & 0 \\
0 & 5 & -4 \\
0 & 0 & 2\sqrt{2}
\end{pmatrix}
\]

The final \( QR \) decomposition is

\[
A = QR \quad Q = \begin{pmatrix}
1/2 & -1/2 & 1/\sqrt{2} \\
1/2 & 1/2 & 0 \\
1/2 & 1/2 & 0 \\
1/2 & -1/2 & -1/\sqrt{2}
\end{pmatrix} \quad R = \begin{pmatrix}
2 & 3 & 0 \\
0 & 5 & -4 \\
0 & 0 & 2\sqrt{2}
\end{pmatrix}
\]
Subspaces

Definition
A **subspace** of $\mathbb{R}^n$ is a subset $V$ of $\mathbb{R}^n$ satisfying:

1. The zero vector is in $V$. **“not empty”**
2. If $u$ and $v$ are in $V$, then $u + v$ is also in $V$. **“closed under addition”**
3. If $u$ is in $V$ and $c$ is in $\mathbb{R}$, then $cu$ is in $V$. **“closed under $\times$ scalars”**

Examples:
- Any span.
- The *column space* of a matrix: $\text{Col } A = \text{Span}\{\text{columns of } A\}$.
- The range of a linear transformation (same as above).
- The *null space* of a matrix: $\text{Nul } A = \{x \mid Ax = 0\}$.
- The *row space* of a matrix: $\text{Row } A = \text{Span}\{\text{rows of } A\}$.
- The $\lambda$-eigenspace of a matrix, where $\lambda$ is an eigenvalue.
- The orthogonal complement $W^\perp$ of a subspace $W$.
- The zero subspace $\{0\}$. 
Definition
Let $V$ be a subspace of $\mathbb{R}^n$. A basis of $V$ is a set of vectors $\{v_1, v_2, \ldots, v_m\}$ in $\mathbb{R}^n$ such that:

1. $V = \text{Span}\{v_1, v_2, \ldots, v_m\}$, and
2. $\{v_1, v_2, \ldots, v_m\}$ is linearly independent.

The number of vectors in a basis is the dimension of $V$, and is written $\text{dim } V$.

Every subspace has a basis, so every subspace is a span. But subspaces have many different bases, and some might be better than others. For instance, Gram–Schmidt takes a basis and produces an orthogonal basis. Or, diagonalization produces a basis of eigenvectors of a matrix.

How do I know if a subset $V$ is a subspace or not?

- Can you write $V$ as one of the examples on the previous slide?
- If not, does it satisfy the three defining properties?

Note on subspaces versus subsets: A subset of $\mathbb{R}^n$ is any collection of vectors whatsoever. Like, the unit circle in $\mathbb{R}^2$, or all vectors with whole-number coefficients. A subspace is a subset that satisfies three additional properties. Most subsets are not subspaces.
Definition
Two $n \times n$ matrices $A$ and $B$ are similar if there is an invertible $n \times n$ matrix $P$ such that

$$A = PBP^{-1}.$$ 

Important Facts:
1. Similar matrices have the same characteristic polynomial.
2. It follows that similar matrices have the same eigenvalues.
3. If $A$ is similar to $B$ and $B$ is similar to $C$, then $A$ is similar to $C$.

Caveats:
1. Matrices with the same characteristic polynomial need not be similar.
2. Similarity has nothing to do with row equivalence.
3. Similar matrices generally do not have the same eigenvectors.
Similarity
Geometric meaning

Let \( A = PBP^{-1} \), and let \( v_1, v_2, \ldots, v_n \) be the columns of \( P \). These form a basis \( B \) for \( \mathbb{R}^n \) because \( P \) is invertible. Key relation: for any vectors \( x, y \) in \( \mathbb{R}^n \),

\[
Ax = y \quad \iff \quad B[x]_B = [y]_B.
\]

This says:

\( A \) acts on the usual coordinates of \( x \) in the same way that \( B \) acts on the \( B \)-coordinates of \( x \).

Example:

\[
A = \begin{pmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix} \quad P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.
\]

Then \( A = PBP^{-1} \). \( B \) acts on the usual coordinates by scaling the first coordinate by 2, and the second by 1/2:

\[
B \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 \\ x_2/2 \end{pmatrix}.
\]

The unit coordinate vectors are eigenvectors: \( e_1 \) has eigenvalue 2, and \( e_2 \) has eigenvalue 1/2.
Examples

Similarity

\[
A = \begin{pmatrix}
\frac{5}{4} & \frac{3}{4} \\
\frac{3}{4} & \frac{5}{4}
\end{pmatrix}
\quad B = \begin{pmatrix}
2 & 0 \\
0 & \frac{1}{2}
\end{pmatrix}
\quad P = \begin{pmatrix}
1 & 1 \\
1 & -1
\end{pmatrix}.
\]

In this case, \(B = \{\frac{1}{1}, \frac{1}{-1}\}\). Let \(v_1 = \frac{1}{1}\) and \(v_2 = \frac{1}{-1}\).

To compute \(y = Ax\):

1. Find \([x]_B\).
2. \([y]_B = B[x]_B\).
3. Compute \(y\) from \([y]_B\).

Say \(x = \begin{pmatrix} 2 \\ 0 \end{pmatrix}\).

1. \(x = v_1 + v_2\) so \([x]_B = \frac{1}{1}\).
2. \([y]_B = B\frac{1}{1} = \begin{pmatrix} 2 \\ \frac{1}{2} \end{pmatrix}\).
3. \(y = 2v_1 + \frac{1}{2}v_2 = \begin{pmatrix} 5/2 \\ 3/2 \end{pmatrix} = Ax\).

Picture:

\(A\) scales the \(v_1\)-coordinate by 2, and the \(v_2\)-coordinate by \(\frac{1}{2}\).
Definition
A matrix equation \( Ax = b \) is **consistent** if it has a solution, and **inconsistent** otherwise.

If \( A \) has columns \( v_1, v_2, \ldots, v_n \), then

\[
\begin{align*}
  b &= Ax = \begin{pmatrix}
  v_1 & v_2 & \cdots & v_m
  \end{pmatrix}
  \begin{pmatrix}
    x_1 \\
    \vdots \\
    x_n
  \end{pmatrix}
  = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n.
\end{align*}
\]

So if \( Ax = b \) has a solution, then \( b \) is a linear combination of \( v_1, v_2, \ldots, v_n \), and conversely. Equivalently, \( b \) is in \( \text{Span}\{v_1, v_2, \ldots, v_n\} = \text{Col} \ A \).

**Important**

\( Ax = b \) is consistent if and only if \( b \) is in \( \text{Col} \ A \).
Suppose that $Ax = b$ is inconsistent. Let $\hat{b} = \text{proj}_{\text{Col } A}(b)$ be the closest vector for which $A\hat{x} = \hat{b}$ does have a solution.

**Definition**

A solution to $A\hat{x} = \hat{b}$ is a **least squares solution** to $Ax = b$. This is the solution $\hat{x}$ for which $A\hat{x}$ is closest to $b$ (with respect to the usual notion of distance in $\mathbb{R}^n$).

**Theorem**

The least-squares solutions to $Ax = b$ are the solutions to

$$A^T A\hat{x} = A^T b.$$ 

If $A$ has *orthogonal* columns $u_1, u_2, \ldots, u_n$, then the least-squares solution is

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

because

$$A\hat{x} = \hat{b} = \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2} u_2 + \cdots + \frac{x \cdot u_m}{u_m \cdot u_m} u_m.$$