Review for Midterm 3

- Bring a **number 2 pencil** to the exam!
- **Extra help session**: today and tomorrow, **4–7pm**, in AH 441
- Room assignments for Thursday, Nov 20, 7-8:15pm:
  - if your last name starts with A-E: 213 Greg Hall
  - if your last name starts with F-L: 100 Greg Hall
  - if your last name starts with M-Sh: 66 Library
  - if your last name starts with Si-Z: 103 Mumford Hall
- Big topics:
  - Orthogonal projections
  - Least squares
  - Gram–Schmidt
  - Determinants
  - Eigenvalues and eigenvectors

Orthogonal projections

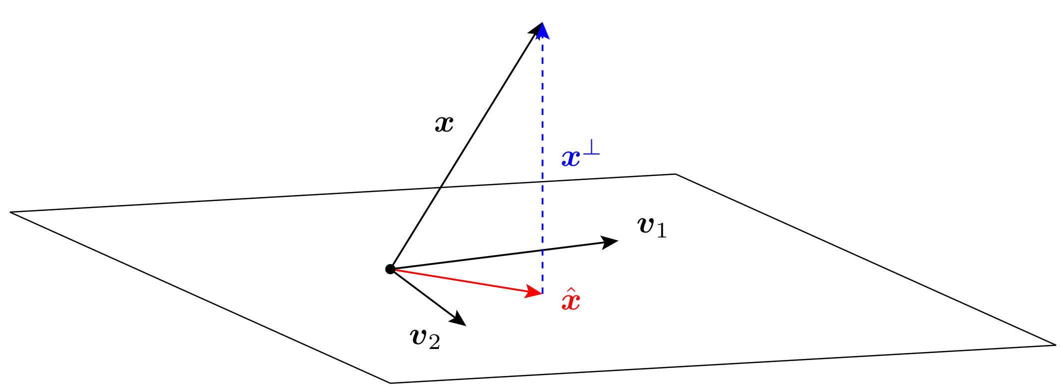
- If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ is an **orthogonal basis** of V, and $\boldsymbol{x}$ is in V, then

$$\boldsymbol{x} = c_1\boldsymbol{v}_1 + \ldots + c_n\boldsymbol{v}_n \qquad \text{with} \quad c_j = \frac{\langle \boldsymbol{x}, \boldsymbol{v}_j \rangle}{\langle \boldsymbol{v}_j, \boldsymbol{v}_j \rangle}.$$

- Suppose that V is a subspace of W, and $\boldsymbol{x}$ is in W, then the **orthogonal projection** of $\boldsymbol{x}$ onto V is given by

$$\hat{\boldsymbol{x}} = c_1\boldsymbol{v}_1 + \ldots + c_n\boldsymbol{v}_n \qquad \text{with} \quad c_j = \frac{\langle \boldsymbol{x}, \boldsymbol{v}_j \rangle}{\langle \boldsymbol{v}_j, \boldsymbol{v}_j \rangle}.$$

  - The basis $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ has to be orthogonal for this formula!!
  - This decomposes $\boldsymbol{x} = \underbrace{\hat{\boldsymbol{x}}}_{\text{in } V} + \underbrace{\boldsymbol{x}^{\perp}}_{\text{in } V^{\perp}}$, where the error $\boldsymbol{x}^{\perp}$ is orthogonal to V. (this decomposition is unique)



Armin Straub
astraub@illinois.edu

- The corresponding **projection matrix** represents $\boldsymbol{x} \mapsto \hat{\boldsymbol{x}}$ with respect to the standard basis.

Example 1.

(a) What is the orthogonal projection of $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ onto $\operatorname{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$?

Solution: The projection is $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

(b) What is the orthogonal projection of $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ onto $\operatorname{span}\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$?

Solution: The projection is $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

Wrong approach!! $\frac{\left\langle \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\rangle}{\left\langle \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\rangle} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \frac{\left\langle \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\rangle}{\left\langle \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\rangle} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

This is wrong because $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ are not orthogonal. (See next example!)

(c) What is the orthogonal projection of $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ onto $\operatorname{span}\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$?

Solution: The projection is $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$.

Wrong!! $\frac{\left\langle \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\rangle}{\left\langle \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\rangle} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \frac{\left\langle \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\rangle}{\left\langle \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\rangle} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \frac{2}{3} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

Corrected: $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ (for instance, using Gram–Schmidt)

$$\frac{\left\langle \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\rangle}{\left\langle \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\rangle} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \frac{\left\langle \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\rangle}{\left\langle \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\rangle} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

(d) What is the projection matrix corresponding to orthogonal projection onto $\operatorname{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$?

Solution: The projection matrix is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

What would Gram–Schmidt do? $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

(e) What is the orthogonal projection of $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ onto $\operatorname{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$?

Armin Straub
astraub@illinois.edu

Solution: The projection is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

- The space of all nice functions with period 2π has the natural inner product $\langle f, g\rangle = \int_0^{2\pi} f(x)g(x)\mathrm{d}x$. [in $\mathbb{R}^n$: $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = x_1y_1 + \ldots + x_ny_n$]
- The functions

$$1, \cos(x), \sin(x), \cos(2x), \sin(2x), \ldots$$

  are an orthogonal basis for this space.
- Expanding a function $f(x)$ in this basis produces its **Fourier series**

$$f(x) = a_0 + a_1\cos(x) + b_1\sin(x) + a_2\cos(2x) + b_2\sin(2x) + \cdots$$

Example 2. How can we compute b_2?

Solution.

$b_2\sin(2x)$ is the orthogonal projection of f onto the span of $\sin(2x)$.

Hence:

$$b_2 = \frac{\langle f(x), \sin(2x)\rangle}{\langle \sin(2x), \sin(2x)\rangle} = \frac{\int_0^{2\pi} f(x)\sin(2x)\mathrm{d}x}{\int_0^{2\pi} \sin^2(2x)\mathrm{d}x}$$

Least squares

- $\hat{\boldsymbol{x}}$ is a **least squares solution** of the system $A\boldsymbol{x} = \boldsymbol{b}$.

  $\Longleftrightarrow$ $\hat{\boldsymbol{x}}$ is such that $A\hat{\boldsymbol{x}} - \boldsymbol{b}$ is as small as possible.

  $\Longleftrightarrow$ $A^TA\hat{\boldsymbol{x}} = A^T\boldsymbol{b}$ (the **normal equations**)

Example 3. Find the least squares line for the data points $(2,1), (5,2), (7,3), (8,3)$.

Solution.

Looking for β_1, β_2 such that the line $y = \beta_1 + \beta_2 x$ best fits the data.

The equations $y_i = \beta_1 + \beta_2 x_i$ in matrix form:

$$\underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{bmatrix}}_{\text{design matrix } X} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}}_{\text{observation vector } \boldsymbol{y}}$$

Here, we need to find a least squares solution to

$$\begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}.$$

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$X^T \boldsymbol{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

Solving $\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \hat{\beta} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$, we find $\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}$.

Gram–Schmidt

Recipe. (Gram–Schmidt orthonormalization)

Given a basis $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$, produce an orthonormal basis $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$.

$$\begin{aligned} \boldsymbol{b}_1 &= \boldsymbol{a}_1, & \boldsymbol{q}_1 &= \frac{\boldsymbol{b}_1}{\|\boldsymbol{b}_1\|} \\ \boldsymbol{b}_2 &= \boldsymbol{a}_2 - \langle \boldsymbol{a}_2, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1, & \boldsymbol{q}_2 &= \frac{\boldsymbol{b}_2}{\|\boldsymbol{b}_2\|} \\ \boldsymbol{b}_3 &= \boldsymbol{a}_3 - \langle \boldsymbol{a}_3, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1 - \langle \boldsymbol{a}_3, \boldsymbol{q}_2 \rangle \boldsymbol{q}_2, & \boldsymbol{q}_3 &= \frac{\boldsymbol{b}_3}{\|\boldsymbol{b}_3\|} \\ &\vdots \end{aligned}$$

- An **orthogonal matrix** is a square matrix Q with orthonormal columns. Equivalently, $Q^T Q = I$ (also true for non-square matrices).
- Apply Gram–Schmidt to the (independent) columns of A to obtain the **QR decomposition** $A = QR$.
  - Q has orthonormal columns (the output vectors of Gram–Schmidt)
  - $R = Q^T A$ is upper triangular

Example 4. Find the QR decomposition of $A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 0 & 1 \end{bmatrix}$.

Solution. We apply Gram–Schmidt to the columns of A:

$$\frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \boldsymbol{q}_1$$

$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \left\langle \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \boldsymbol{q}_1 \right\rangle \boldsymbol{q}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{\sqrt{5}}\frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 4/5 \\ -2/5 \\ 1 \end{bmatrix}, \qquad \frac{1}{\sqrt{9/5}}\begin{bmatrix} 4/5 \\ -2/5 \\ 1 \end{bmatrix} = \boldsymbol{q}_2$$

Hence: $Q = [\ \boldsymbol{q}_1\ \ \boldsymbol{q}_2\ ] = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{4}{\sqrt{45}} \\ \frac{2}{\sqrt{5}} & -\frac{2}{\sqrt{45}} \\ 0 & \frac{5}{\sqrt{45}} \end{bmatrix}$

And: $R = Q^T A = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \\ \frac{4}{\sqrt{45}} & -\frac{2}{\sqrt{45}} & \frac{5}{\sqrt{45}} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{5}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ 0 & \frac{9}{\sqrt{45}} \end{bmatrix}$

Determinants

- A is invertible $\iff$ $\det(A) \neq 0$
- $\det(AB) = \det(A)\det(B)$
- $\det(A^{-1}) = \frac{1}{\det(A)}$
- $\det(A^T) = \det(A)$
- The **determinant** is characterized by:
  - the normalization $\det I = 1$,
  - and how it is affected by elementary row operations:
    - **(replacement)** Add one row to a multiple of another row.

      Does not change the determinant.
    - **(interchange)** Interchange two rows.

      Reverses the sign of the determinant.
    - **(scaling)** Multiply all entries in a row by s.

      Multiplies the determinant by s.

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 1 & 5 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 3 & 5 \end{vmatrix} \overset{R4 \to R4 - \frac{3}{2}R3}{=} \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 1 & 5 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & \frac{7}{2} \end{vmatrix} = 1 \cdot 2 \cdot 2 \cdot \frac{7}{2} = 14$$

Armin Straub
astraub@illinois.edu

- **Cofactor expansion** is another way to compute determinants.

$$\begin{vmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ 2 & 0 & 1 \end{vmatrix} = -2 \cdot \begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} + (-1) \cdot \begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & 0 \\ 3 & 2 \end{vmatrix}$$

$$= -2 \cdot (-1) + (-1) \cdot 1 - 0 = 1$$

Example 5. What is $\begin{vmatrix} 1 & 1 & 1 & 4 \\ -1 & 2 & 2 & 5 \\ 0 & 3 & 3 & 1 \\ 2 & 0 & 0 & 5 \end{vmatrix}$?

Solution. The determinant is 0 because the matrix is not invertible (second and third column are the same).

Eigenvalues and eigenvectors

- If $A\boldsymbol{x} = \lambda \boldsymbol{x}$, then $\boldsymbol{x}$ is an **eigenvector** of A with **eigenvalue** λ.
- λ is an eigenvalue of A $\iff$ $\underbrace{\det(A - \lambda I)}_{\text{characteristic polynomial}} = 0$.

  Why? Because $A\boldsymbol{x} = \lambda \boldsymbol{x} \iff (A - \lambda I)\boldsymbol{x} = \boldsymbol{0}$.
- The **eigenspace** of λ is $\mathrm{Nul}(A - \lambda I)$.

  It consists of all eigenvectors (plus $\boldsymbol{0}$) with eigenvalue λ.
- Eigenvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ of A corresponding to different eigenvalues are independent.
- Useful for checking: sum of eigenvalues $=$ sum of diagonal entries