Announcements Dec 1

87.77

76.02

- Masks ⇝ Thank you!
- CIOS ⇝ Quiz drop! 85% by Dec 6
- Remainings WeBWorKs not for a grade, for practice only
- Office hrs: Tue 4-5 Teams + Thu 1-2 Skiles courtyard/Teams + Pop-ups
- Cumulative Final exam Tue Dec 14 6-8:50 pm on Teams.

- Many TA office hours listed on Canvas
- PLUS sessions: Tue 6–7 GT Connector, Thu 6–7 BlueJeans
- Indoor Math Lab: Mon–Thu 11–6, Fri 11–3 Clough 246 + 252
- Outdoor Math Lab: Tue–Thu 2–4 Skiles Courtyard
- Virtual Math Lab `https://tutoring.gatech.edu/drop-in/`
- Section M web site: Google "Dan Margalit math", click on 1553
  - future blank slides, past lecture slides, advice
- Old exams: Google "Dan Margalit math", click on Teaching
- Tutoring: `http://tutoring.gatech.edu/tutoring`
- Counseling center: `https://counseling.gatech.edu`
- Use Piazza for general questions
- You can do it!

Section 6.3

Orthogonal projection

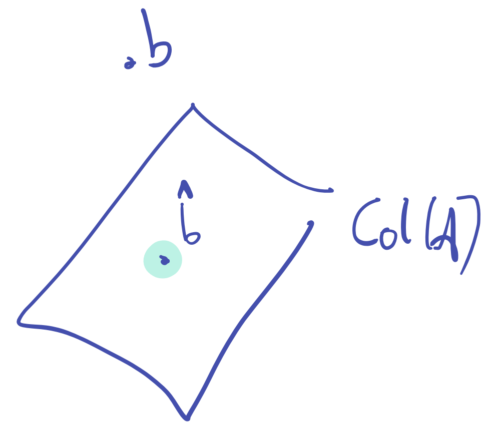
Orthogonal Projections

Theorem. Let $W = \mathrm{Col}(A)$. For any vector b in $\mathbb{R}^n$, the equation

$$A^T A x = A^T b$$

is consistent and the orthogonal projection b_W is equal to Ax where x is any solution.

$\hat{b} = b_W$



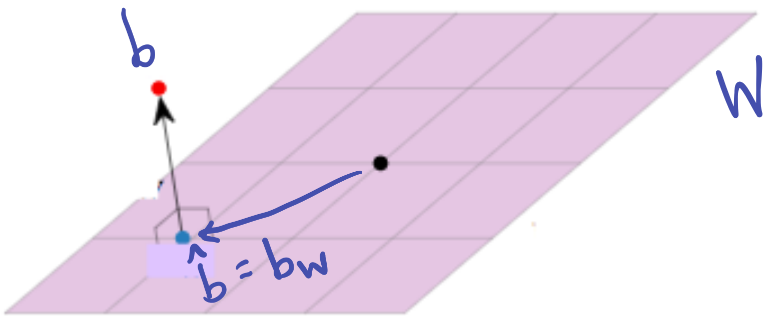
$\hat{b} = A\hat{x}$

where $\hat{x}$ is a soln to $A^T A x = A^T b$

Orthogonal Projections

Let b be a vector in $\mathbb{R}^n$ and W a subspace of $\mathbb{R}^n$.

The orthogonal projection of b onto W the vector obtained by drawing a line segment from b to W that is perpendicular to W.



Fact. The following three things are all the same:

- The orthogonal projection of b onto W
- The vector b_W (the W-part of b) algebra!
- The closest vector in W to b geometry!

Orthogonal Projections

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Example. Find b_W ($= \hat{b}$) if $b = \begin{pmatrix} 6 \\ 5 \\ 4 \end{pmatrix}$, $W = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$

Steps. Find A^TA and A^Tb, then solve for x, then compute Ax.

① $A^TA = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

② $A^Tb = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 6 \\ 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 10 \\ 11 \end{pmatrix}$

③ $\left(\begin{array}{cc|c} 2 & 1 & 10 \\ 1 & 2 & 11 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|c} 1 & 2 & 11 \\ 2 & 1 & 10 \end{array}\right) \rightsquigarrow \begin{pmatrix} 1 & 2 & 11 \\ 0 & -3 & -12 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 11 \\ 0 & 1 & 4 \end{pmatrix} \rightsquigarrow \left(\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 4 \end{array}\right)$

$\left(A^TA \mid A^Tb \right)$

④ $\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 7 \\ 4 \\ 3 \end{pmatrix}$ ← b_W

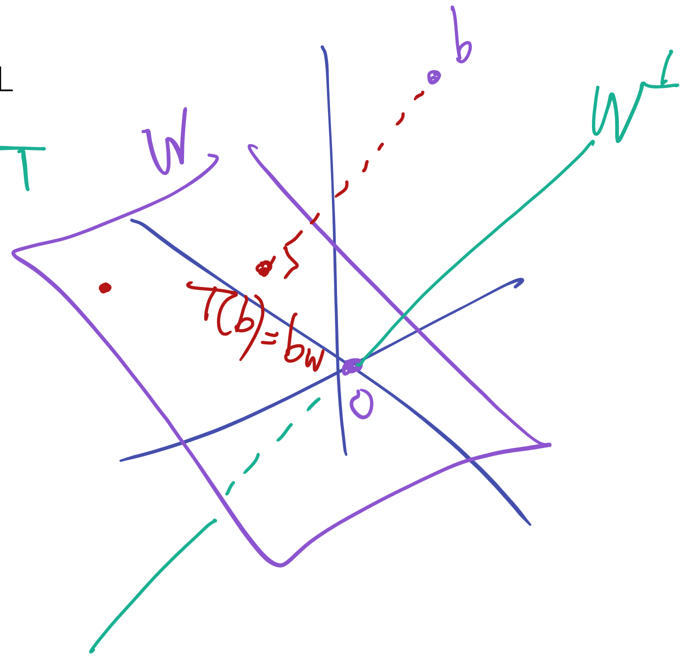
Question. How far is b from W?

$$\|b_{W^\perp}\| = \|b - b_W\| = \left\| \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\| = \sqrt{(-1)^2 + 1^2 + 1^2} = \sqrt{3}$$

Projections as linear transformations

Let W be a subspace of $\mathbb{R}^n$ and let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the function given by $T(b) = b_W$ (orthogonal projection). Then

- T is a linear transformation
- $T(b) = b$ if and only if b is in W
- $T(b) = 0$ if and only if b is in $W^\perp$
- $T \circ T = T$ $T \circ T \circ T \circ T \circ T \circ T = T$
- The range of T is W



Properties of projection matrices

Let W be a subspace of $\mathbb{R}^n$ and let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the function given by $T(b) = b_W$ (orthogonal projection). Let A be the standard matrix for T. Then

- The 1–eigenspace of A is W (unless $W = 0$)
- The 0–eigenspace of A is $W^\perp$ (unless $W = \mathbb{R}^n$)
- $A^2 = A$
- $\mathrm{Col}(A) = W$
- $\mathrm{Nul}(A) = W^\perp$
- A is diagonalizable; its diagonal matrix has m 1's & $n - m$ 0's where $m = \dim W$ (this gives another way to find A)

diag'able: sum of dims of eigensp's is n.
and: $\dim W + \dim W^\perp = n$.

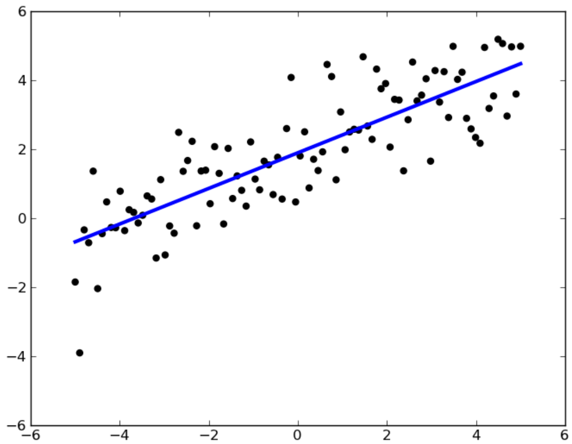
You can check these properties for the matrix in the last example. It would be very hard to prove these facts without any theory. But they are all easy once you know about linear transformations!

Section 6.5

Least Squares Problems

Least Squares problems

What if we can't solve $Ax = b$? How can we solve it as closely as possible?



To solve $Ax = b$ as closely as possible, we orthogonally project b onto $\mathrm{Col}(A)$; call the result $\widehat{b}$. Then solve $Ax = \widehat{b}$. This is the *least squares solution* to $Ax = b$.

Outline of Section 6.5

- The method of least squares
- Application to best fit lines/planes
- Application to best fit curves

Least squares solutions

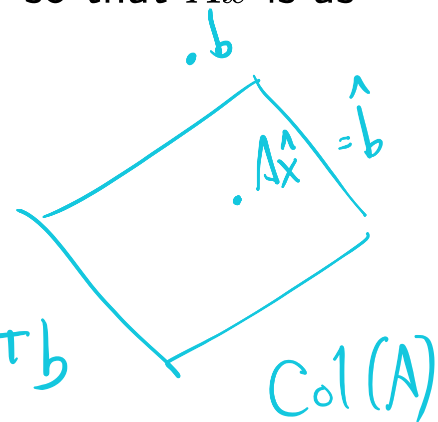
$A = m \times n$ matrix.

A least squares solution to $Ax = b$ is an $\widehat{x}$ in $\mathbb{R}^n$ so that $A\widehat{x}$ is as close as possible to b, this means $A\hat{x} = \hat{b}$.

The error is $\|A\widehat{x} - b\|$.

In last section we wanted $\hat{b}$.

Now we want $\hat{x}$.



Solve $A^TAx = A^Tb$

(like before)

but don't multiply the answer by A.

▸ Demo

Least squares solutions

A least squares solution to $Ax = b$ is an $\widehat{x}$ in $\mathbb{R}^n$ so that $A\widehat{x}$ is as close as possible to b. $\|b_{W^\perp}\|$

The error is $\|A\widehat{x} - b\|$. = distance from $\hat{b}$ to b.

Theorem. The least squares solutions to $Ax = b$ are the solutions to

$$(A^T A)x = (A^T b)$$

So this is just like what we did before when we were finding the projection of b onto $\mathrm{Col}(A)$. But now we just solve and don't (necessarily) multiply the solution by A.

Least squares solutions

Example

Theorem. The least squares solutions to $Ax = b$ are the solutions to

$$(A^T A)x = (A^T b)$$

Find the least squares solutions to $Ax = b$ for this A and b:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$$

What is the error?

Least squares solutions

Example

Formula: $(A^T A)x = (A^T b)$

Find the least squares solution/error to $Ax = b$:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$$

$b - \hat{b} = b_{W^\perp}$

① $A^T A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ 3 & 3 \end{pmatrix}$

② $A^T b = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \end{pmatrix}$

③ $\left(\begin{array}{cc|c} 5 & 3 & 0 \\ 3 & 3 & 6 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|c} 1 & 1 & 2 \\ 5 & 3 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & -2 & -10 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & 5 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 1 & 0 & -3 \\ 0 & 1 & 5 \end{array}\right)$

④ Answer: $\hat{x} = \begin{pmatrix} -3 \\ 5 \end{pmatrix}$ (Don't multiply by A)

⑤ $\hat{b} = A\hat{x} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} -3 \\ 5 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ -1 \end{pmatrix}$

$\|b - \hat{b}\| = \left\| \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\|$

$= \sqrt{1^2 + (-2)^2 + 1^2}$

$= \sqrt{6}$ ← error.

Least squares solutions

Theorem. Let A be an $m \times n$ matrix. The following are equivalent:

1. $Ax = b$ has a unique least squares solution for all b in $\mathbb{R}^n$
2. The columns of A are linearly independent
3. $A^T A$ is invertible

In this case the least squares solution is $(A^T A)^{-1}(A^T b)$.

Application

Best fit lines

Eqn of line: $y = Mx + B$. Want M, B.

Problem. Find the best-fit line through $(0,6)$, $(1,0)$, and $(2,0)$.

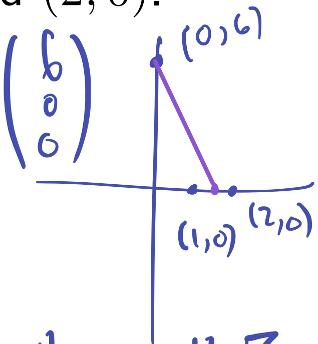
$(0,6)$: $6 = M\cdot 0 + B \cdot 1$

$(1,0)$: $0 = M\cdot 1 + B \cdot 1$

$(2,0)$: $0 = M\cdot 2 + B \cdot 1$

$\rightsquigarrow A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix}$ $b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$

$x = \begin{pmatrix} M \\ B \end{pmatrix}$



No soln! No line going thru all 3.

Demo

We already found the least squares soln:

$$\hat{x} = \begin{pmatrix} -3 \\ 5 \end{pmatrix} = \begin{pmatrix} M \\ B \end{pmatrix}$$

Best fit line: $y = -3x + 5$.

Best fit lines

Poll

What does the best fit line minimize?

1. the sum of the squares of the distances from the data points to the line
2. the sum of the squares of the vertical distances from the data points to the line
3. the sum of the squares of the horizontal distances from the data points to the line
4. the maximal distance from the data points to the line

Least Squares Problems

More applications

Determine the least squares problem $Ax = b$ to find the best parabola $y = Cx^2 + Dx + E$ for the points:

$$(0,0), (2,0), (3,0), (0,1)$$

plug in pts ⇝ 4 eqns in
3 vars (C, D, E).

Demo

Least Squares Problems

More applications

Determine the least squares problem $Ax = b$ to find the best fit ellipse $Cx^2 + Dxy + Ey^2 + Fx + Gy + H = 0$ for the points:

$$(0, 0), (2, 0), (3, 0), (0, 1)$$

Gauss invented the method of least squares to predict the orbit of the asteroid Ceres as it passed behind the sun in 1801.

Least Squares Problems

Best fit plane

Determine the least squares problem $Ax = b$ to find the best fit linear function $f(x, y) = Cx + Dy + E$

x	y	$f(x,y)$
1	0	0
0	1	1
-1	0	3
0	-1	4

Summary of Section 6.5

- A least squares solution to $Ax = b$ is an $\widehat{x}$ in $\mathbb{R}^n$ so that $A\widehat{x}$ is as close as possible to b.
- The error is $\|A\widehat{x} - b\|$.
- The least squares solutions to $Ax = b$ are the solutions to $(A^T A)x = (A^T b)$.
- To find a best fit line/parabola/etc. write the general form of the line/parabola/etc. with unknown coefficients and plug in the given points to get a system of linear equations in the unknown coefficients.

Typical Exam Questions 6.5

- Find the best fit line through $(1,0)$, $(2,1)$, and $(3,1)$. What is the error?
- Find the best fit parabola through $(1,0)$, $(2,1)$, $(3,1)$, and $(3,0)$. What is the error?
- True/false. For every set of three points in $\mathbb{R}^2$ there is a unique best fit line.
- True/false. If $\widehat{x}$ is the least squares solution to $Ax = b$ for an $m \times n$ matrix A, then $\widehat{x}$ is the closest point in $\mathbb{R}^n$ to b.
- True/false. If $\widehat{x}$ and $\widehat{y}$ are both least squares solutions to $Ax = b$ then $\widehat{x} - \widehat{y}$ is in the null space of A.