Announcements April 20

- Final exam Mon Apr 27 6 am – Tue Apr 28 11:59 pm.
- Final exam open for 3 hrs, planned to be a 1 hour exam.
- CIOS-like survey on Canvas Quizzes
- Chapter 6 WeBWorKs are not graded.
- My office hours Monday 3-4, Wed 2-3, and by appointment
- Stay tuned for review sessions...
- TA office hours on Blue Jeans (you can go to any of these!)
  - Isabella Wed 11-12
  - Kyle Wed 3-5, Thu 1-3
  - Kalen Mon/Wed 1-2
  - Sidhanth Tue 10-12
- Supplemental problems & practice exams on master web site
- Counseling Center: http://counseling.gatech.edu Click

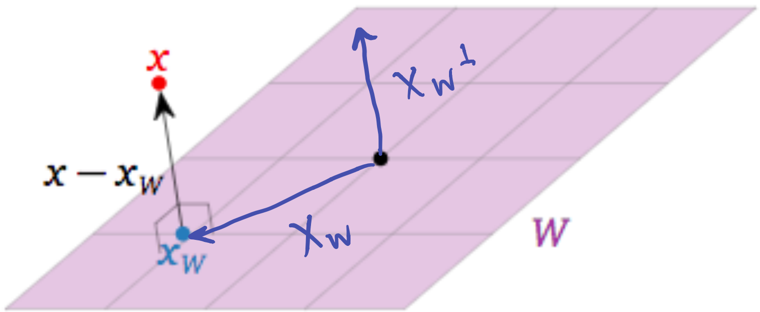
Section 6.3

Orthogonal projection

Orthogonal Projections

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

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



Fact. The following three things are all the same:

- The orthogonal projection of v onto W
- The vector v_W (the W-part of v)
- The closest vector in W to v

Orthogonal Projections

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

$$A^T A x = A^T v$$

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

Why? Choose $\widehat{x}$ so that $A\widehat{x} = v_W$. We know $v - v_W = v - A\widehat{x}$ is in $W^\perp = \mathrm{Nul}(A^T)$ and so

$$0 = A^T(v - A\widehat{x}) = A^T v - A^T A\widehat{x}$$

$$\rightsquigarrow A^T A\widehat{x} = A^T v$$

Point: Being $\perp$ to Col (A) $\iff$ being in Nul A^T

geometry

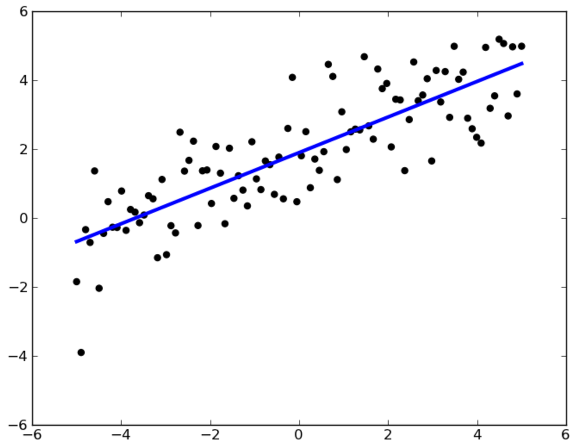
algebra

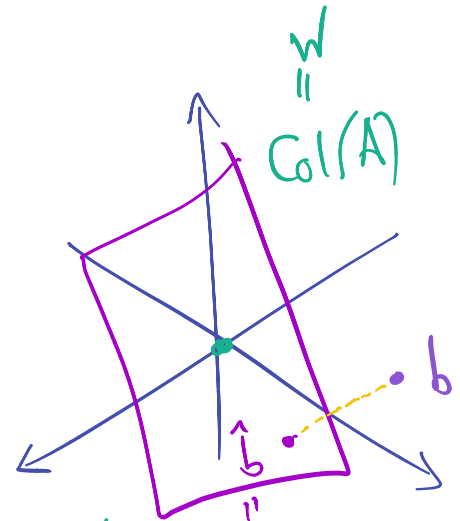
Section 6.5

Least Squares Problems

Least Squares problems

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





b_W = projection

$\rightsquigarrow$ Solve $Ax = \hat{b}$

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

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 $Ax = b$.

looking for

Outline of Section 6.5

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

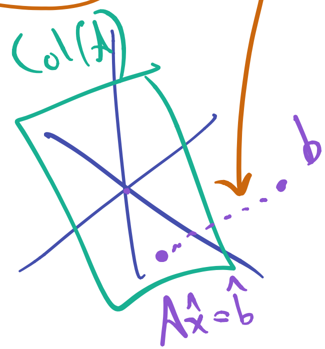
Least squares solutions

$A = m \times n$ 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.

$= \hat{b}$

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



The $A\hat{x}$'s are all vectors in... Col(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.

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

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:

$$A = \begin{array}{c} \begin{array}{cc} -6 & \;\;6 \end{array} \\ \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \end{array} \qquad b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$$

Is there an actual solution? (consistent?)

No.

What is the error?

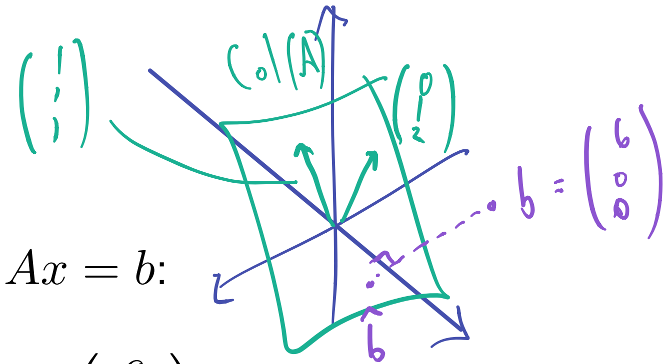
Least squares solutions

Example

Formula: $(A^TA)x = (A^Tb)$

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

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



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

$$A^TA = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ 3 & 3 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \end{pmatrix}$$

setup

$$\begin{pmatrix} 5 & 3 \\ 3 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \end{pmatrix} \sim \left(\begin{array}{cc|c} 5 & 3 & 0 \\ 3 & 3 & 6 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 1 & 2 \\ 5 & 3 & 0 \end{array}\right)$$

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

row red.

$$\sim \begin{array}{l} x_1 = -3 \\ x_2 = 5 \end{array} \qquad \hat{x} = \begin{pmatrix} -3 \\ 5 \end{pmatrix}$$

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

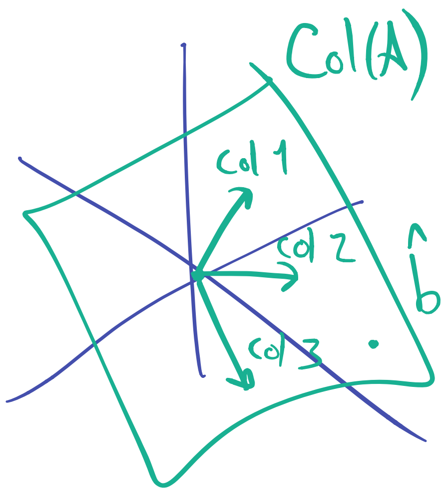
Error: $\|b - \hat{b}\| = \left\| \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 5 \\ 2 \\ -1 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\| = \sqrt{6}$

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 $(A^T A)^{-1}(A^T b)$.



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

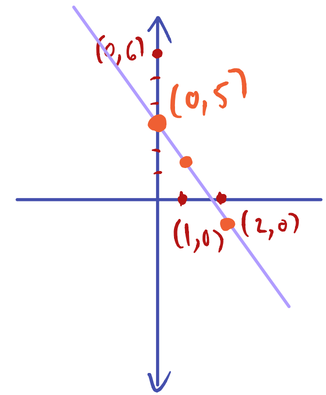
unique soln $\iff$ A^TA invertible.

Application

Best fit lines

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

Want $y = Mx + B$ (as close as possible)



Given points $\rightsquigarrow$

$$6 = M\cdot 0 + 1\cdot B$$
$$0 = M\cdot 1 + 1\cdot B$$
$$0 = M\cdot 2 + 1\cdot B$$

Demo

$$\rightsquigarrow \begin{array}{c} \begin{array}{cc} M & B \end{array} \\ \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \end{array} x = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$$

y-int: 6, 2, 4, (5

slope: -2, -4, -3, -1

$$\xrightarrow{\text{as above}} \hat{x} = \begin{pmatrix} -3 \\ 5 \end{pmatrix}$$

$$\rightsquigarrow M = -3,\ B = 5 \rightsquigarrow y = -3x + 5$$

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

y-vals of orig pts

$$\left\| \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 5 \\ 2 \\ -1 \end{pmatrix} \right\| = \sqrt{(6-5)^2 + (0-2)^2 + (0-(-1))^2}$$

Least Squares Problems

More applications

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

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

$(0,0)$: $0 = 0 \cdot C + 0 \cdot D + 1 \cdot E$

$(2,0)$: $0 = 4 \cdot C + 2 \cdot D + 1 \cdot E$

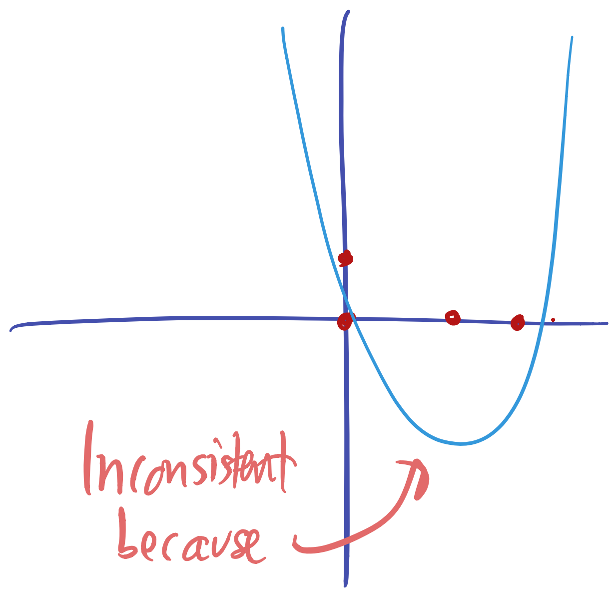
$(3,0)$: $0 = 9 \cdot C + 3 \cdot D + 1 \cdot E$

$(0,1)$: $1 = 0 \cdot C + 0 \cdot D + 1 \cdot E$

▸ Demo

$$\begin{pmatrix} 0 & 0 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} C \\ D \\ E \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Inconsistent because

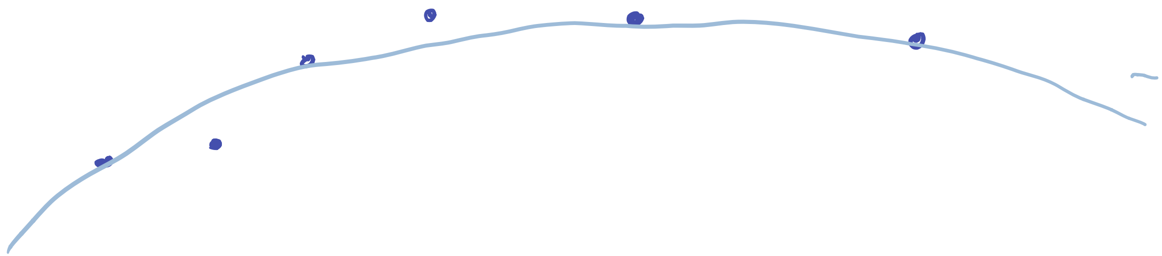


Least Squares Problems

More applications

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

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



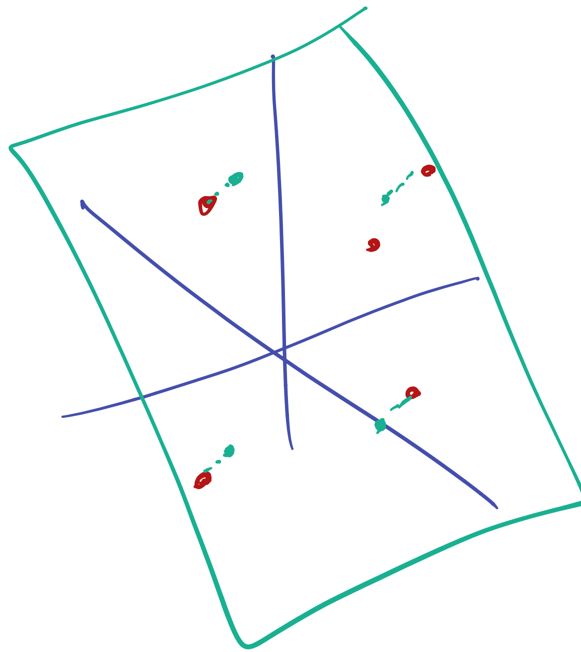
Gauss invented the method of least squares to predict the orbit of the asteroid Ceres as it passed behind the sun in 1801.

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



$0 = 1 \cdot C + 0 \cdot D + 1 \cdot E$

$1 = 0 \cdot C + 1 \cdot D + 1 \cdot E$

$3 = -1 \cdot C + 0 \cdot D + 1 \cdot E$

etc. . .

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 $(A^TA)x = (A^Tb)$.
- 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.