

## Announcements

Wednesday, November 28

- ▶ Please fill out your CIOS survey!

If 85% of the class completes the survey by 11:59pm on December 7, then we will drop two quizzes instead of one.

- ▶ Final exam time: **Tuesday, December 11, 6–8:50pm.**
- ▶ WeBWorK 6.6, 7.1, 7.2 are due today.
- ▶ No quiz on Friday! But this is the only recitation on chapter 7.
- ▶ My office is Skiles 244 and Rabinoffice hours are: Mondays, 12–1pm; Wednesdays, 1–3pm.

## Section 7.5

### The Method of Least Squares

## Motivation

We now are in a position to solve the motivating problem of this third part of the course:

**Problem**

Suppose that  $Ax = b$  does not have a solution. What is the best possible approximate solution?

To say  $Ax = b$  does not have a solution means that  $b$  is not in  $\text{Col } A$ .

The closest possible  $\hat{b}$  for which  $Ax = \hat{b}$  does have a solution is  $\hat{b} = b_{\text{Col } A}$ .

Then  $A\hat{x} = \hat{b}$  is a consistent equation.

A solution  $\hat{x}$  to  $A\hat{x} = \hat{b}$  is a **least squares solution**.

# Least Squares Solutions

Let  $A$  be an  $m \times n$  matrix.

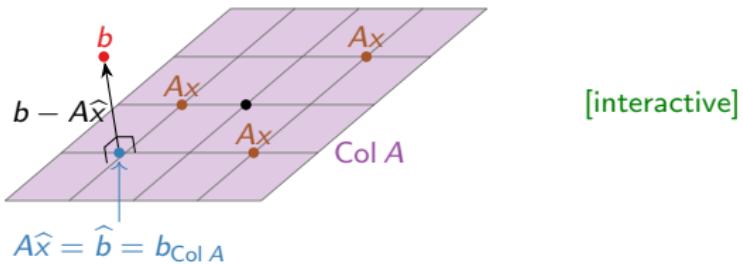
## Definition

A **least squares solution** of  $Ax = b$  is a vector  $\hat{x}$  in  $\mathbb{R}^n$  such that

$$\|b - A\hat{x}\| \leq \|b - Ax\|$$

for all  $x$  in  $\mathbb{R}^n$ .

Note that  $b - A\hat{x}$  is in  $(\text{Col } A)^\perp$ .



In other words, a least squares solution  $\hat{x}$  solves  $Ax = b$  as closely as possible.

Equivalently, a least squares solution to  $Ax = b$  is a vector  $\hat{x}$  in  $\mathbb{R}^n$  such that

$$A\hat{x} = \hat{b} = b_{\text{Col } A}.$$

This is because  $\hat{b}$  is the closest vector to  $b$  such that  $A\hat{x} = \hat{b}$  is consistent.

## Least Squares Solutions

### Computation

We want to solve  $A\hat{x} = \hat{b} = b_{\text{Col } A}$ . Or,  $A\hat{x} = b_W$  for  $W = \text{Col } A$ .

To compute  $b_W$  we need to solve  $A^T A v = A^T b$ ; then  $b_W = A v$ .

**Conclusion:**  $\hat{x}$  is just a solution of  $A^T A v = A^T b$ !

### Theorem

The least squares solutions of  $Ax = b$  are the solutions of

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

Note we compute  $\hat{x}$  directly, without computing  $\hat{b}$  first.

## Least Squares Solutions

### Example

Find the least squares solutions of  $Ax = b$  where:

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

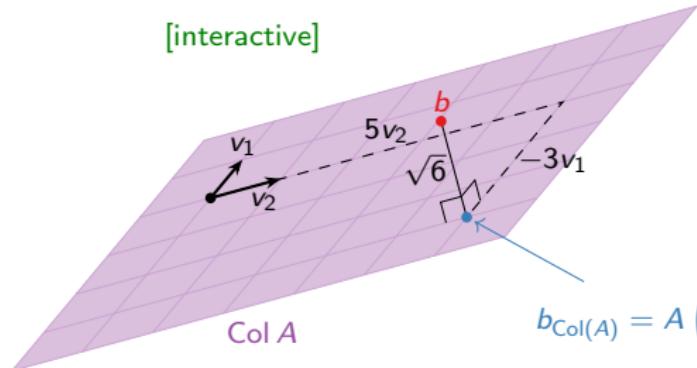
So the only least squares solution is  $\hat{x} = \begin{pmatrix} -3 \\ 5 \end{pmatrix}$ .

# Least Squares Solutions

Example, continued

How close did we get?

[interactive]



$$b_{\text{Col}(A)} = A \begin{pmatrix} -3 \\ 5 \end{pmatrix}$$

Note that

$$\begin{pmatrix} -3 \\ 5 \end{pmatrix}$$

records the coefficients of  $v_1$  and  $v_2$  in  $\hat{b}$ .

## Least Squares Solutions

### Second example

Find the least squares solutions of  $Ax = b$  where:

$$A = \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{pmatrix} \quad b = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

So the only least squares solution is  $\hat{x} = \begin{pmatrix} 1/3 \\ -1/3 \end{pmatrix}$ .

[\[interactive\]](#)

# Least Squares Solutions

## Uniqueness

When does  $Ax = b$  have a *unique* least squares solution  $\hat{x}$ ?

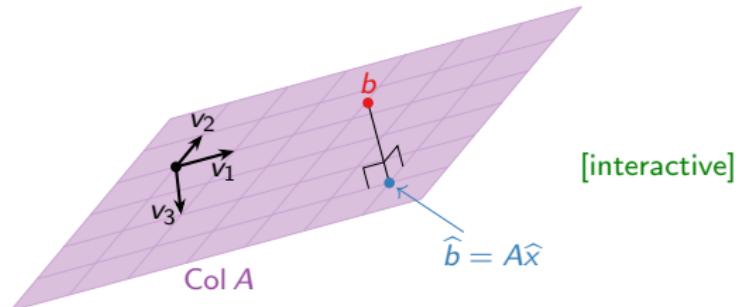
### Theorem

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  $\mathbb{R}^m$ .
2. The columns of  $A$  are linearly independent.
3.  $A^T A$  is invertible.

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

**Why?** If the columns of  $A$  are linearly *dependent*, then  $A\hat{x} = \hat{b}$  has many solutions:



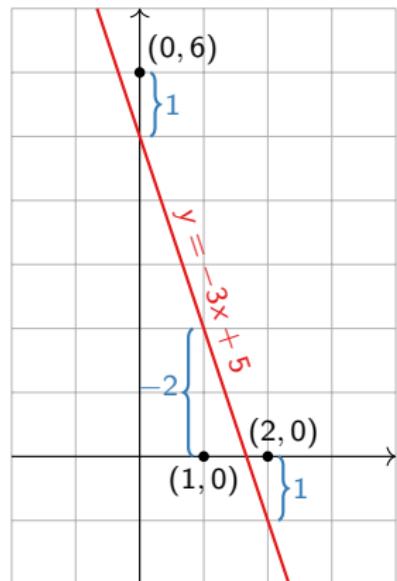
**Note:**  $A^T A$  is always a square matrix, but it need not be invertible.

# Application

Data modeling: best fit line

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

[interactive]



$$A \begin{pmatrix} 5 \\ -3 \end{pmatrix} - \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

Poll

## Application

### Best fit ellipse

Find the best fit ellipse for the points  $(0, 2)$ ,  $(2, 1)$ ,  $(1, -1)$ ,  $(-1, -2)$ ,  $(-3, 1)$ ,  $(-1, -1)$ .

The general equation for an ellipse is

$$x^2 + Ay^2 + Bxy + Cx + Dy + E = 0$$

So we want to solve:

$$\begin{aligned} (0)^2 + A(2)^2 + B(0)(2) + C(0) + D(2) + E &= 0 \\ (2)^2 + A(1)^2 + B(2)(1) + C(2) + D(1) + E &= 0 \\ (1)^2 + A(-1)^2 + B(1)(-1) + C(1) + D(-1) + E &= 0 \\ (-1)^2 + A(-2)^2 + B(-1)(-2) + C(-1) + D(-2) + E &= 0 \\ (-3)^2 + A(1)^2 + B(-3)(1) + C(-3) + D(1) + E &= 0 \\ (-1)^2 + A(-1)^2 + B(-1)(-1) + C(-1) + D(-1) + E &= 0 \end{aligned}$$

In matrix form:

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

## Application

Best fit ellipse, continued

$$A = \begin{pmatrix} 4 & 0 & 0 & 2 & 1 \\ 1 & 2 & 2 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 4 & 2 & -1 & -2 & 1 \\ 1 & -3 & -3 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 0 \\ -4 \\ -1 \\ -1 \\ -9 \\ -1 \end{pmatrix}.$$

$$A^T A = \begin{pmatrix} 36 & 7 & -5 & 0 & 12 \\ 7 & 19 & 9 & -5 & 1 \\ -5 & 9 & 16 & 1 & -2 \\ 0 & -5 & 1 & 12 & 0 \\ 12 & 1 & -2 & 0 & 6 \end{pmatrix} \quad A^T b = \begin{pmatrix} -19 \\ 17 \\ 20 \\ -9 \\ -16 \end{pmatrix}$$

Row reduce:

$$\left( \begin{array}{ccccc|c} 36 & 7 & -5 & 0 & 12 & -19 \\ 7 & 19 & 9 & -5 & 1 & 17 \\ -5 & 9 & 16 & 1 & -2 & 20 \\ 0 & -5 & 1 & 12 & 0 & -9 \\ 12 & 1 & -2 & 0 & 6 & -16 \end{array} \right) \xrightarrow{\text{~~~~~}} \left( \begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & 405/266 \\ 0 & 1 & 0 & 0 & 0 & -89/133 \\ 0 & 0 & 1 & 0 & 0 & 201/133 \\ 0 & 0 & 0 & 1 & 0 & -123/266 \\ 0 & 0 & 0 & 0 & 1 & -687/133 \end{array} \right)$$

Best fit ellipse:

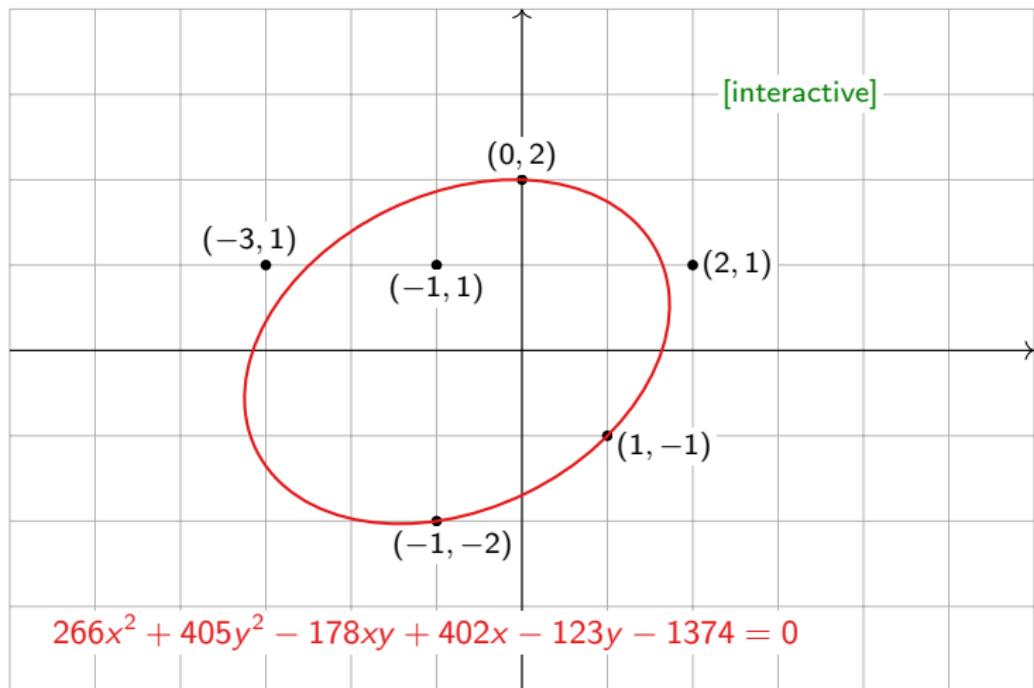
$$x^2 + \frac{405}{266}y^2 - \frac{89}{133}xy + \frac{201}{133}x - \frac{123}{266}y - \frac{687}{133} = 0$$

or

$$266x^2 + 405y^2 - 178xy + 402x - 123y - 1374 = 0.$$

# Application

Best fit ellipse, picture



**Remark:** Gauss invented the method of least squares to do exactly this: he predicted the (elliptical) orbit of the asteroid Ceres as it passed behind the sun in 1801.

## Application

### Best fit parabola

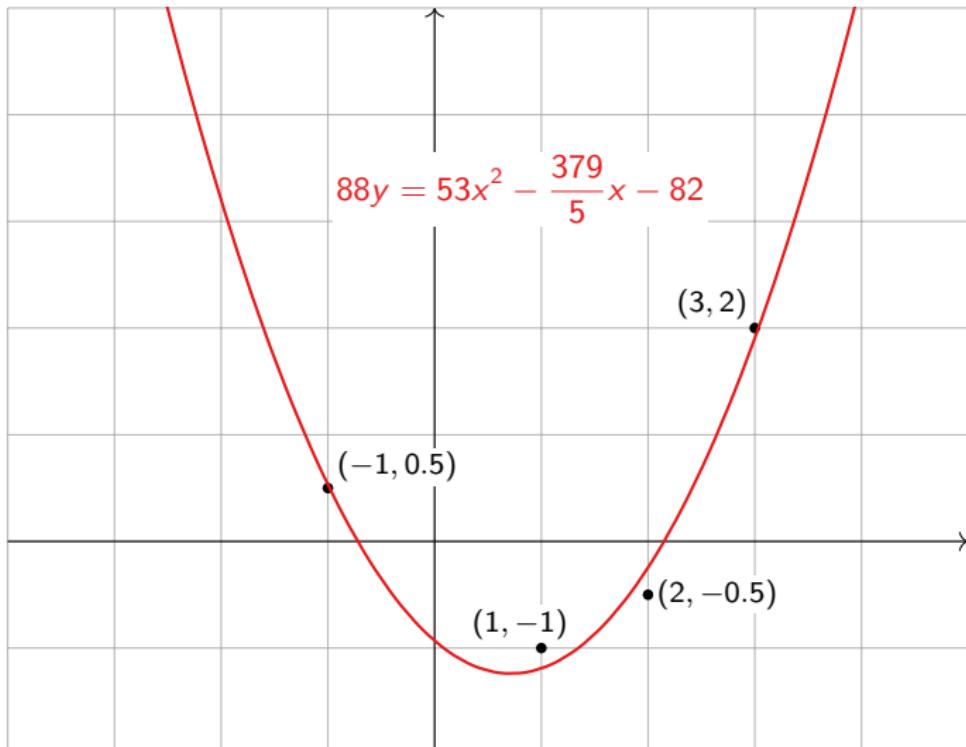
What least squares problem  $Ax = b$  finds the best parabola through the points  $(-1, 0.5), (1, -1), (2, -0.5), (3, 2)$ ?

Answer:

$$88y = 53x^2 - \frac{379}{5}x - 82$$

# Application

Best fit parabola, picture



[interactive]

## Application

### Best fit linear function

What least squares problem  $Ax = b$  finds the best linear function  $f(x, y)$  fitting the following data?

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

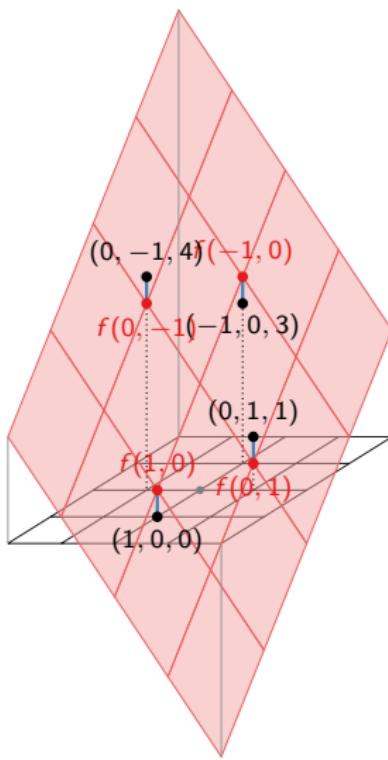
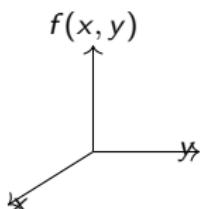
Answer:

$$f(x, y) = -\frac{3}{2}x - \frac{3}{2}y + 2$$

# Application

Best fit linear function, picture

[interactive]



Graph of

$$f(x, y) = -\frac{3}{2}x - \frac{3}{2}y + 2$$

# Application

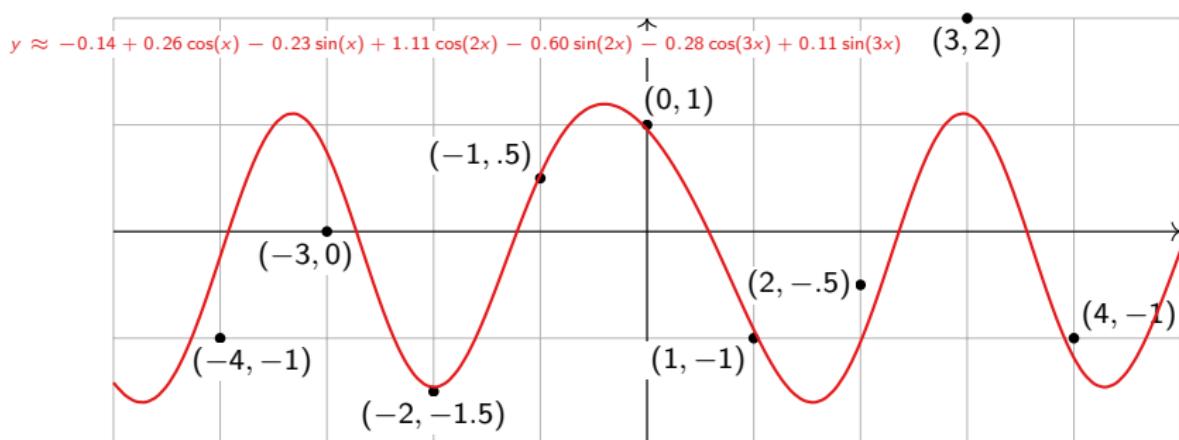
## Best-fit Trigonometric Function

For fun: what is the best-fit function of the form

$$y = A + B \cos(x) + C \sin(x) + D \cos(2x) + E \sin(2x) + F \cos(3x) + G \sin(3x)$$

passing through the points

$$\begin{pmatrix} -4 \\ -1 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ -1.5 \end{pmatrix}, \begin{pmatrix} -1 \\ .5 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ -.5 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 4 \\ -1 \end{pmatrix}?$$



[interactive]

## Summary

- ▶ A **least squares solution** of  $Ax = b$  is a vector  $\hat{x}$  such that  $\hat{b} = A\hat{x}$  is as close to  $b$  as possible.
- ▶ This means that  $\hat{b} = b_{\text{Col } A}$ .
- ▶ One way to compute a least squares solution is by solving the system of equations

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

Note that  $A^T A$  is a (symmetric) square matrix.

- ▶ Least-squares solutions are unique when the columns of  $A$  are linearly independent.
- ▶ You can use least-squares to find best-fit lines, parabolas, ellipses, planes, etc.