~~HW5 will be posted before Tues class, will be due before Fri class.~~

For the rest of this course we will discuss Two Major Applications of linear algebra:

1. Least Squares Approximation.
2. Diagonalization (or "Spectral Analysis")

First: Least Squares Approximation.

Suppose that the linear system

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

has no solution. In this case we would like to find some $\vec{x}$ such that the distance/length $\|A\vec{x} - \vec{b}\|$ is as small as possible.

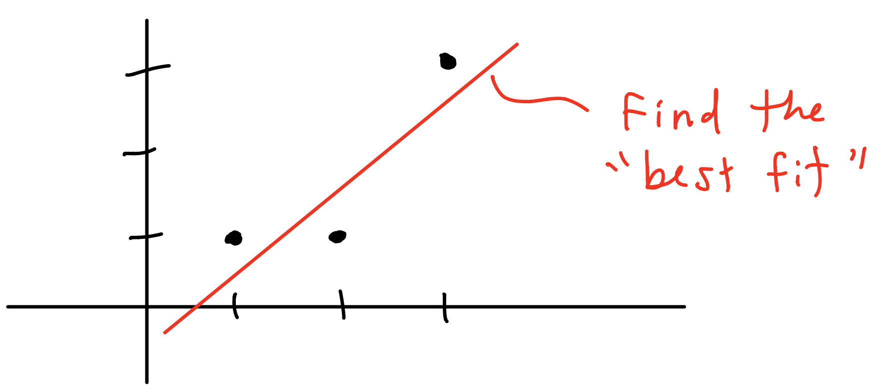
Under good conditions (namely, when the columns of A are independent) then it turns out that the solution of this problem is easy to compute:

$$\vec{x} = (A^T A)^{-1} A^T \vec{b}.$$

Before I tell you why this works, let's see an example.

Problem: Find the line $y = b + cx$ that is "closest" to the points

$$(x,y) = (1,1), (2,1), (3,3).$$



There are different ways to interpret the words "closest" & "best fit." First I will show you the standard answer.

Put the problem in Matrix notation:

Optimistically, if the points actually lie on a line $y = b + cx$ then we have three linear equations

data point (x, y)	equation $y = b + cx$
$(1, 1)$	$1 = b + c$
$(2, 1)$	$1 = b + 2c$
$(3, 3)$	$3 = b + 3c$

We can express these as a single matrix equation:

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$$

Does this have an exact solution?

Recall: 3 equations in 2 unknowns probably has NO SOLUTION!

According to what I said above, we have an <u>approximate</u> solution given by

$$\begin{pmatrix} b \\ c \end{pmatrix} = \left[\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \right]^{-1} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$$

$$\begin{pmatrix} b \\ c \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}.$$

PAUSE. To compute the inverse of a 2x2 matrix, there is a shortcut:

$$\begin{pmatrix} e & f \\ g & h \end{pmatrix}^{-1} = \frac{1}{eh - fg} \begin{pmatrix} h & -f \\ -g & e \end{pmatrix}.$$

Hence, $\begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix}^{-1} = \frac{1}{6} \begin{pmatrix} 14 & -6 \\ -6 & 3 \end{pmatrix}.$

UNPAUSE.

$$\begin{pmatrix} b \\ c \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 14 & -6 \\ -6 & 3 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$$

$$= \frac{1}{6}\begin{pmatrix} 14 & -6 \\ -6 & 3 \end{pmatrix}\begin{pmatrix} 5 \\ 12 \end{pmatrix}$$

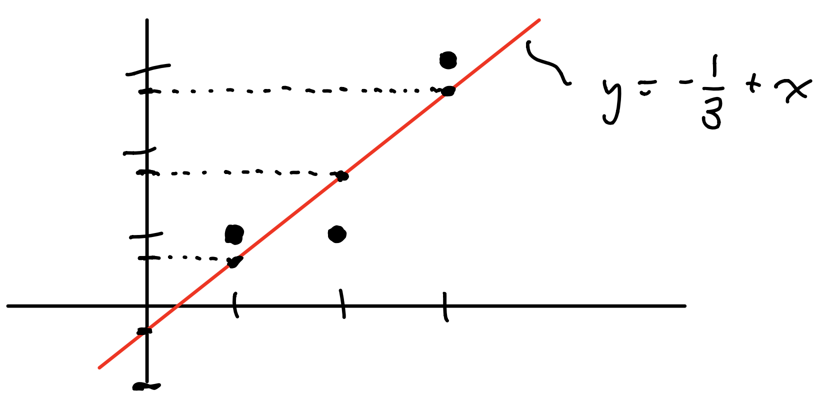
$$= \frac{1}{6}\begin{pmatrix} -2 \\ 6 \end{pmatrix}$$

$$= \begin{pmatrix} -1/3 \\ 1 \end{pmatrix}$$

Conclusion: The best fit line is

$y = b + cx$

$y = -\frac{1}{3} + x.$



In what sense is this a "best fit"?

Recall: I claim that the approximation

$$\vec{x} = (A^T A)^{-1} A^T \vec{b}$$

minimizes the length $\|A\vec{x} - \vec{b}\|$,

hence also minimizes the squared length:

$$\|A\vec{x} - \vec{b}\|^2.$$

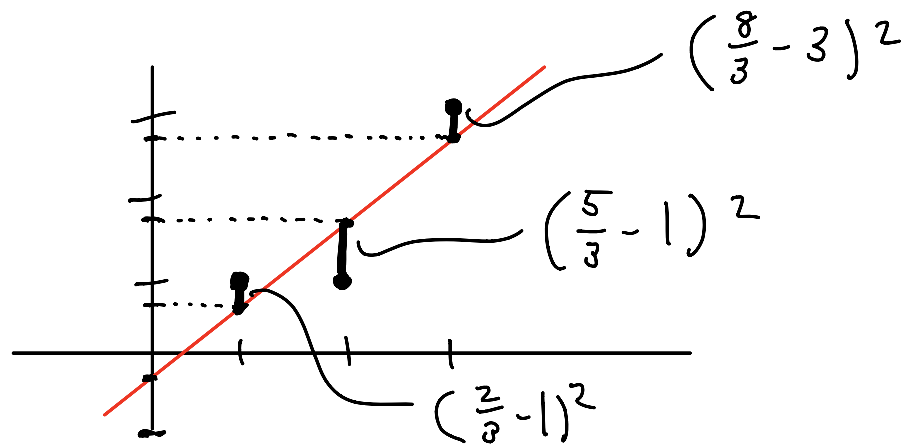
In our example, we have minimized

$$\left\| \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} -1/3 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \right\|^2$$

$$= \left\| \begin{pmatrix} 2/3 \\ 5/3 \\ 8/3 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \right\|^2$$

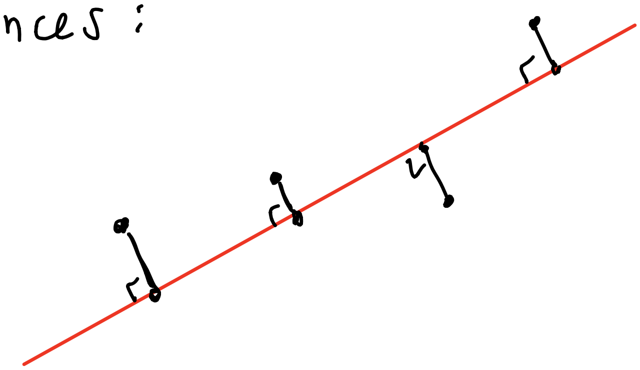
$$= \left(\frac{2}{3} - 1\right)^2 + \left(\frac{5}{3} - 1\right)^2 + \left(\frac{8}{3} - 3\right)^2.$$

Geometrically, this is the sum of the squares of the vertical errors:



This is called the OLS ("ordinary least squares") regression line.

You might prefer to minimize the sum of (the squares of) the orthogonal distances:



This is called the TLS ("total least squares") regression line, which is much harder to compute.

[Remark : The TLS approximation of $A\vec{x} = \vec{b}$ is given by

$$\vec{x} = (A^T A - \sqrt{\lambda} I)^{-1} A^T \vec{b},$$

where λ is the "smallest eigenvalue" of the matrix $A^T A$.]

Why does it work?

The column space of an $m \times n$ matrix

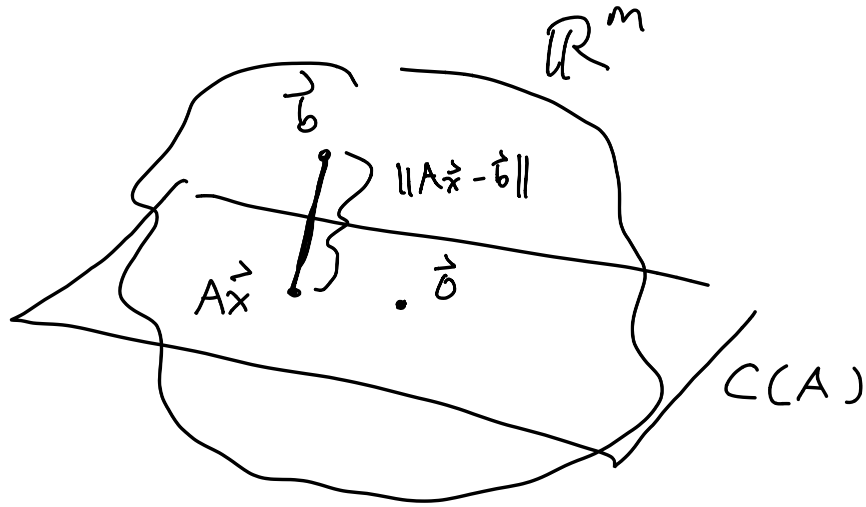
$$A = \left(\vec{a}_1 \; \vec{a}_2 \cdots \vec{a}_n \right)$$

consists of all vectors of the form

$$A\vec{x} = x_1 \vec{a}_1 + x_2 \vec{a}_2 + \cdots + x_n \vec{a}_n.$$

To say that $A\vec{x} = \vec{b}$ has <u>no solution</u> means that the point $\vec{b}$

is <u>not</u> in the column space of A:



we want to find some point $A\vec{x}$ in the column space such that the distance $\|A\vec{x}-\vec{b}\|$ is minimized.

This is a geometric problem:

"Given a subspace $U \subseteq \mathbb{R}^m$ and a point $\vec{b} \in \mathbb{R}^m$ <u>not</u> in the subspace, Find the point $\vec{u} \in U$ that is <u>closest</u> to $\vec{b}$."

We will solve this next time!