

Getting close ('least squares')

Math 4A – Scharlemann

11 March 2015



CELL PHONES OFF

Final: Here 8 a.m. Monday. Do only **10 (ten)** out of 12 problems.

Must bring

- TA.RD.IS code (on Gauchospace)
- Seat assignment (on Gauchospace - same as last time)
- Convincing photo id
- 8.5" by 11" bluebook for **clear written work**.

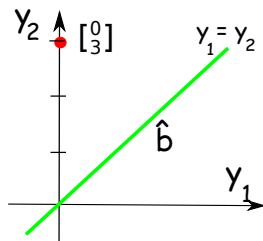
May bring

- **One** 3" x 5" card
- Headlamp

Don't bring

- Calculator
- Telephone or other electronics

iClicker question



Let $\hat{b}$ be the point on the line $y_1 = y_2$ at which the line gets closest to $\begin{bmatrix} 0 \\ 3 \end{bmatrix}$, i. e. the point $y_1 = 0, y_2 = 3$. What is $\hat{b}$?

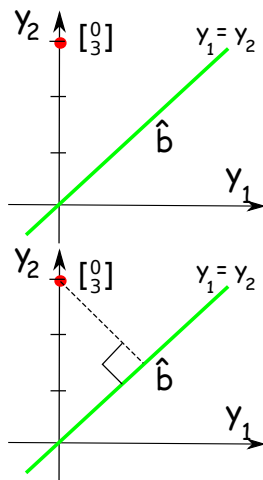
A) $\hat{b} = \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \end{bmatrix}$

B) $\hat{b} = \begin{bmatrix} 3/2 \\ 3/2 \end{bmatrix}$

C) $\hat{b} = \begin{bmatrix} 2/3 \\ 2/3 \end{bmatrix}$

D) $\hat{b} = \begin{bmatrix} \sqrt{3} \\ \sqrt{3} \end{bmatrix}$

Discussion



Geometers tell us: to find the closest point on a line to a given point, **drop a perpendicular** from the point to the line.

Symmetry says that height of $\hat{b}$ is half that of $\begin{bmatrix} 0 \\ 3 \end{bmatrix}$ so $\hat{b} = \begin{bmatrix} ? \\ 3/2 \end{bmatrix}$.
 $\hat{b}$ is on line $y_1 = y_2$ so $= \begin{bmatrix} 3/2 \\ 3/2 \end{bmatrix}$.

Dot product verifies: $\left(\begin{bmatrix} 0 \\ 3 \end{bmatrix} - \vec{\hat{b}} \right) \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/2 \\ 3/2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{3}{2} - \frac{3}{2} = 0$

Suppose you are given a system of linear equations that is **inconsistent**:

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m$$

Inconsistent $\implies$ there is **no solution**.

Even if we can't hit the values $b_1, \dots, b_m$ with any choice of $x_1, \dots, x_n$, how close can we get to $b_1, \dots, b_m$?

And what, exactly, does this question even mean?

Translation to vector thinking:

Given: an $m \times n$ matrix A and a vector equation $A\vec{x} = \vec{b}$.

So far we have learned how to answer four questions:

- Does the vector equation have a solution $\vec{x} \in \mathbb{R}^n$?
- If it does have a solution, is the solution unique?
- If the solution is unique, find it.
- If the solution is not unique, describe all solutions.

Typical answer: $\vec{v}_s + \text{Span}\{\vec{v}_1, \dots, \vec{v}_k\} \subset \mathbb{R}^n$.

But what if there is **no solution**? I.e what if the answer to the first question is “no”?

The situation: $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $\vec{b} \notin \text{Image}(A) \subset \mathbb{R}^m$.

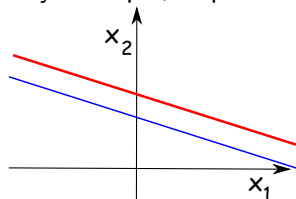
Three part question:

- How close does the subspace $\text{Image}(A) \subset \mathbb{R}^m$ come to the point $\vec{b} \in \mathbb{R}^m$?
- What is the closest point $\vec{\hat{b}}$ in $\text{Image}(A)$ to $\vec{b} \in \mathbb{R}^m$?
- For what values of $\vec{x} \in \mathbb{R}^n$ is $A\vec{x} = \vec{\hat{b}}$? That is, describe all $\vec{x} \in \mathbb{R}^n$ so that $A\vec{x}$ is as close as possible to $\vec{b}$.

Then these values of $\vec{x} \in \mathbb{R}^n$ are in some sense best possible values to use. Note that the idea of “close” here only makes sense now that we have defined distance in $\mathbb{R}^m$.

(The hope: close enough for government work! 😊)

Toy example, in pictures from lecture 1:



$$1x_1 + 2x_2 = 3$$

$$1x_1 + 2x_2 = 4$$

inconsistent

From this picture, no clue how to pick 'best' x_1, x_2 .

The sophisticated Math 4A point of view:

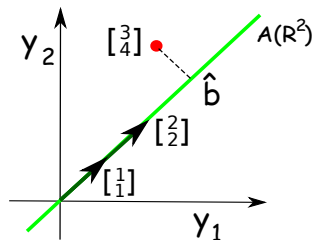
Given a matrix $A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ defining a linear transf. $\mathbb{R}^2 \rightarrow \mathbb{R}^2$.

Try to pick $\vec{x} \in \mathbb{R}^2$ so that $A\vec{x}$ is as close as possible to $\begin{bmatrix} 3 \\ 4 \end{bmatrix} \in \mathbb{R}^2$

$$\vec{b} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \text{ lives.}$$

- $A(\vec{e}_1) = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \mathbb{R}^2$ and
- $A(\vec{e}_2) = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \mathbb{R}^2.$

Hence $\text{Image}(A)$ is the line of all multiples of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ in $\mathbb{R}^2$



Know two things about $\vec{b}$:

- $\vec{b} = A\vec{x}$ for **some** $\vec{x}$ and
- vector $\vec{b} - \hat{\vec{b}} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} - \hat{\vec{b}}$ is $\perp$ to $A\vec{x}$ for **all** $\vec{x}$

Vector $\vec{b} - \hat{\vec{b}} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} - \hat{\vec{b}}$ is $\perp$ to $A\vec{x}$ for **all** $\vec{x}$ means

$A\vec{x} \cdot (\vec{b} - \hat{\vec{b}}) = 0$ for **all** $\vec{x}$.

Now apply alternate view of dot product: $\vec{a} \cdot \vec{b} = (\vec{a}^T)\vec{b}$

Last statement becomes

$$(A\vec{x})^T(\vec{b} - \hat{\vec{b}}) = \vec{x}^T A^T(\vec{b} - \hat{\vec{b}}) = \vec{x} \cdot A^T(\vec{b} - \hat{\vec{b}}) = 0 \text{ for all } \vec{x}$$

Since it's true for all $\vec{x}$, in particular it's true for $\vec{x} = A^T(\vec{b} - \hat{\vec{b}})$:

$$A^T(\vec{b} - \hat{\vec{b}}) \cdot A^T(\vec{b} - \hat{\vec{b}}) = 0 \implies A^T(\vec{b} - \hat{\vec{b}}) = \vec{0} \implies A^T\vec{b} = A^T\hat{\vec{b}}.$$

Substituting $\hat{\vec{b}} = A\vec{x}$ gives the matrix equation

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

Can solve for $\vec{x}$, and then $\hat{\vec{b}} = A\vec{x}$.

In toy example this becomes

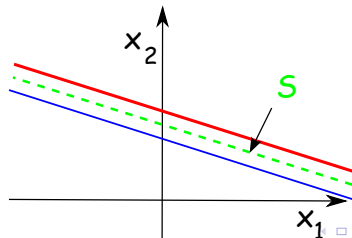
$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Calculation simplifies this to

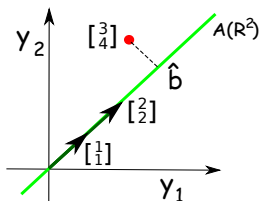
$$\begin{bmatrix} 7 \\ 14 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Solution found by usual techniques

$$S = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix} + \text{Span} \left\{ \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}$$



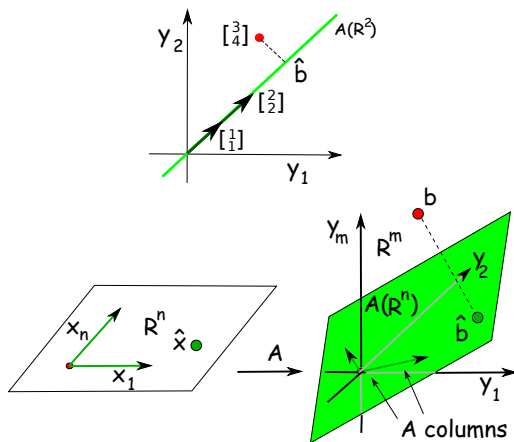
When put in the linear equations, result is as close as we can hope to come to $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$. How close can we come?



$$\vec{\hat{b}} = A\vec{x} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 3.5 \\ 3.5 \end{bmatrix} \Rightarrow |\vec{b} - \vec{\hat{b}}| = \frac{\sqrt{2}}{2}$$

General case is argued just the same – picture more complicated (more dimensions), but same outcome:

Replace $\vec{b} = A\vec{x}$ with $A^T \vec{b} = A^T A \vec{x}$.



Example: Find the solution that gets as close as possible to

$$x_1 - x_2 = 4$$

$$3x_1 + 2x_2 = 1$$

$$-2x_1 + 4x_2 = 3$$

Matrix equation $A\vec{x} = \vec{b}$:

$$\begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

Try for a solution:

$$\begin{bmatrix} 1 & -1 & 4 \\ 3 & 2 & 1 \\ -2 & 4 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -1 & 4 \\ 0 & 5 & -11 \\ 0 & 2 & 11 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -1 & 4 \\ 0 & 5 & -11 \\ 0 & 0 & (\frac{7}{5})11 \end{bmatrix} \Rightarrow \text{Fail!}$$

No solution for the equation

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

so instead seek closest possible by solving

$$A^T A \vec{x} = A^T \vec{b}:$$

$$\begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} \Rightarrow$$
$$\begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

To solve

$$\begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

Find reduced echelon form:

$$\begin{bmatrix} 14 & -3 & | & 1 \\ -3 & 21 & | & 10 \end{bmatrix} \rightarrow \begin{bmatrix} 14 & -3 & | & 1 \\ 0 & \frac{285}{14} & | & \frac{143}{14} \end{bmatrix} \rightarrow \begin{bmatrix} 14 & -3 & | & 1 \\ 0 & 1 & | & \frac{143}{285} \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} 14 & 0 & | & \frac{143+95}{95} \\ 0 & 1 & | & \frac{143}{285} \end{bmatrix} = \begin{bmatrix} 14 & 0 & | & \frac{238}{95} \\ 0 & 1 & | & \frac{143}{285} \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & | & \frac{17}{95} \\ 0 & 1 & | & \frac{143}{285} \end{bmatrix}$$

Hence

$$x_1 = \frac{17}{95}$$

$$x_2 = \frac{143}{285}$$

is best (but not fully correct) solution.

Then calculate

$$\vec{\tilde{b}} = A\vec{x} = \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} \frac{17}{95} \\ \frac{143}{285} \end{bmatrix} = \begin{bmatrix} -\frac{92}{285} \\ \frac{439}{285} \\ \frac{94}{57} \end{bmatrix} \sim \begin{bmatrix} -0.323 \\ 1.54 \\ 1.65 \end{bmatrix}$$

Sadly, not particularly close to $\vec{b} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$.

But we now know that it's the closest we can get!