24. XI. 22

Goal
- Best fit problems
- least squares solutions
- Two ways to find least squares solutions

Today $F = \mathbb{R}$, $\mathbb{R}^n$ carries dot product.

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{A} & \mathbb{R}^m \\ \rotatebox{90}{$\in$} & & \\ \hat{x} & \longmapsto & A\hat{x} \end{array}$$

min's dist from b

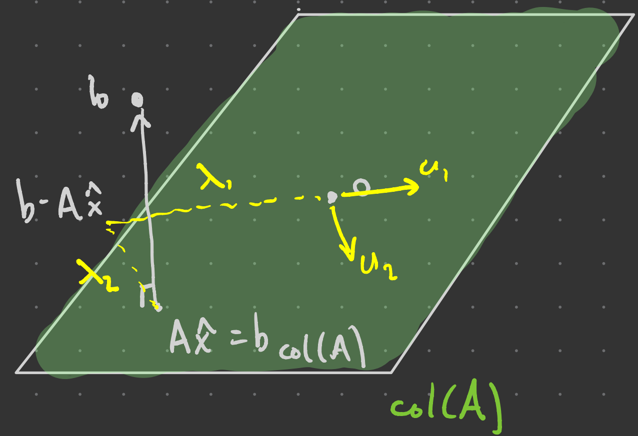
Defn For $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, a least squares solution of $Ax = b$ is a vector $\hat{x} \in \mathbb{R}^n$ such that

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

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

Note $\overset{\operatorname{im} A =}{\{ Ax \mid x \in \mathbb{R}^n \}} = \operatorname{col}(A)$, the column space of A.

So $A\hat{x} = b_{\operatorname{col}(A)}$, the orthogonal proj'n of b onto $\operatorname{col}(A)$.



$A = \begin{pmatrix} | & | \\ u_1 & u_2 \\ | & | \end{pmatrix}$

$A\hat{x} = \lambda_1 u_1 + \lambda_2 u_2$

The entries of $\hat{x}$ are the coordinates $\Rightarrow \mathrm{Rep}_{\{u_1, u_2\}} \hat{x} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$

of $b_{col(A)}$ wrt columns of A when the cols are lin ind.

Thm A vector $\hat{x} \in \mathbb{R}^n$ is the least squares solution of $Ax = b$ iff it's a solution of the associated normal system $A^T A x = A^T b$.

Pf idea Norm of $r(x) = b - Ax$ is minimized when it is orthogonal to the row space of A.

Show that $\text{row}(A)^{\perp} = \ker(A^T)$ so $\hat{x}$ satisfies

$$A^T r(\hat{x}) = 0 \iff A^T(b - A\hat{x}) = 0$$

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

Fact The normal system $A^T A x = A^T b$ has a unique sol'n iff $\ker A = 0$ iff columns of A are lin ind.

Algorithm For $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$

1. Compute $A^T A$, $A^T b$.
2. Row reduce $[A^T A \mid A^T b]$
3. This is always consistent and any sol'n $\hat{x}$ is a least squares sol'n.

<u>Note</u> When $Ax = b$ has a <u>unique</u> least squares sol'n, it is

$$\hat{x} = (\underset{\mathbb{R}^{m\times m}}{A^TA})^{-1}\underset{\mathbb{R}^{m\times n}}{A^T}\, b.$$

Now suppose the colums of A are orthogonal, say

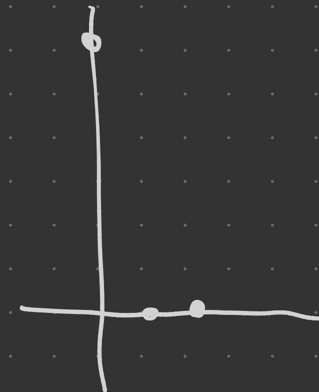
$$A = \begin{pmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_n \\ | & | & & | \end{pmatrix}.$$

Then $\mathrm{col}(A)$ has orthogonal basis $u_1, \ldots, u_n$ and

$$b_{\mathrm{col}(A)} = \frac{\langle b, u_1\rangle}{\|u_1\|^2}u_1 + \cdots + \frac{\langle b, u_n\rangle}{\|u_n\|^2} = A\underbrace{\begin{pmatrix} \langle b, u_1\rangle / \|u_1\|^2 \\ \vdots \\ \langle b, u_n\rangle / \|u_n\|^2 \end{pmatrix}}_{\hat{x}}$$

E.g. (best fit line) Suppose we have points

$$(0,6),\ (1,0),\ (2,0).$$



What line best fits this data?

Eq'n of a line: $\quad y = Mx + B.$

"Want":
$$\begin{aligned} 6 &= M\cdot 0 + B \\ 0 &= M\cdot 1 + B \\ 0 &= M \cdot 2 + B \end{aligned}$$

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

Then we "want" $Ax = b$.

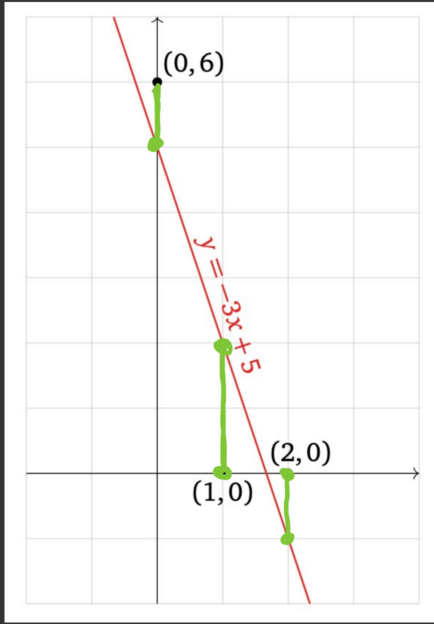
Use the algorithm:

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

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

and $\left[A^T A \mid A^T b\right] = \left(\begin{array}{cc|c} 5 & 3 & 0 \\ 3 & 3 & 6 \end{array}\right) \xrightarrow[G\text{-}J]{} \left(\begin{array}{cc|c} 1 & 0 & -3 \\ 0 & 1 & 5 \end{array}\right)$.

Thus $\hat{x} = \begin{pmatrix} -3 \\ 5 \end{pmatrix}$ is the unique least squares sol'n, and thus the best fit line is $y = -3x + 5$.



Q What has least squares minimized in finding this best fit line?

A $A\hat{x} = \begin{pmatrix} f(0) \\ f(1) \\ f(2) \end{pmatrix}$ and $b - A\hat{x} = \begin{pmatrix} 6 - f(0) \\ 0 - f(1) \\ 0 - f(2) \end{pmatrix}$

$f(t) = \hat{M}x + \hat{B}$

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

so minimizing

$$\|b - A\hat{x}\| = \sqrt{(6 - f(0))^2 + (0 - f(1))^2 + (0 - f(2))^2}$$

$\Updownarrow$

minimizing $\|b - A\hat{x}\|^2 = \cdots$