Lecture 14: Least Squares

Let $A \in M_{m\times n}(\mathbb{C})$, how do you "solve"

$$\underbrace{A\vec{x}}_{\in\mathbb{C}^m} = \underbrace{\vec{b}}_{\in\mathbb{C}^m}$$

the equations are more than likely inconsisistent. Solutions exists if

$$\vec{b} \in \mathrm{im}(A) = \mathrm{span}\{\text{columns of } A\}.$$

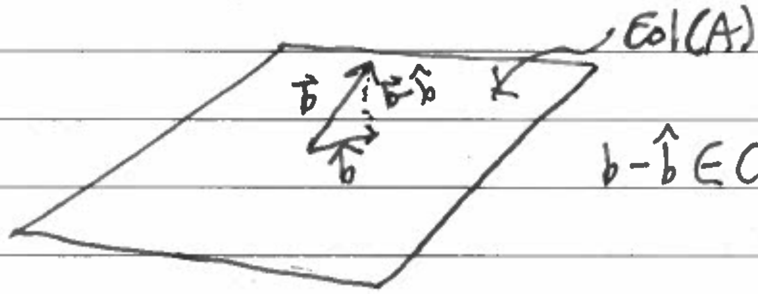
Idea (Gauss): Project b onto columns of A.

$$\Rightarrow A = [\vec{c}_1 | \cdots | \vec{c}_n]$$

$$\Rightarrow \hat{b} = \mathrm{Proj}_{\mathrm{Col}(A)}(\vec{b}) = \langle \vec{b}, \vec{c}_1 \rangle \frac{\vec{c}_1}{\|\vec{c}_1\|} + \ldots + \langle \vec{b}, \vec{c}_n \rangle \frac{\vec{c}_n}{\|\vec{c}_n\|}.$$

Solve instead:

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



$b - \hat{b} \in \mathrm{Col}(A)^\perp$

Why is this a good Idea?

* Really want to find $\vec{x}$ such that $\|b - Ax\|$ is as small as possible.

$$\|b - Ax\|^2 = \|\underbrace{\vec{b} - \hat{b}}_{\in \mathrm{Col}(A)^\perp} + \underbrace{\hat{b} - A\vec{x}}_{\in \mathrm{Col}(A)}\|^2$$

$$= \underbrace{\|\vec{b} - \hat{b}\|^2}_{\#} + \underbrace{\|\hat{b} - A\vec{x}\|^2}_{\text{make as small as possible.}}$$

$\Rightarrow$ Choose $\hat{x}$ such that $A\hat{x} = \hat{b}$!!

$\Rightarrow \|b - A\hat{x}\|^2 = \|\vec{b} - \hat{b}\|$

$\Rightarrow \hat{x}$ is called the least squares solution.

Normal Equations

Recall that

$$\text{Nul}(A^*) = \text{Col}(A)^\perp$$

Since $\vec{b} - \hat{b} \in \text{Col}(A)^\perp$ it follows that

$$A^*(b - \hat{b}) = 0$$

$$\Rightarrow A^* b = A^* \hat{b}$$

$$\Rightarrow \underbrace{A^* A \hat{x} = A^* b}_{\text{Normal Equations}}$$

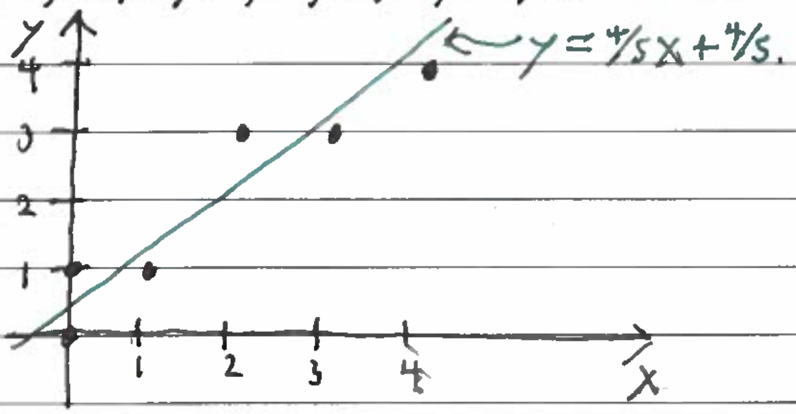
If $A^* A$ is invertible

$$\hat{x} = (A^* A)^{-1} A^* b$$

Example:

Find the least squares line of the data

$(0,1), (1,1), (2,3), (3,3), (4,4)$



Assume $y = ax + b$.

$\Rightarrow$ "Solve"

$$\begin{aligned} 0 \cdot a + b &= 1 \\ 1 \cdot a + b &= 1 \\ 2 \cdot a + b &= 3 \\ 3 \cdot a + b &= 3 \\ 4 \cdot a + b &= 4 \end{aligned} \rightarrow \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 1 \\ 1 \\ 3 \\ 3 \\ 4 \end{bmatrix}}_{\vec{b}}$$

$$\Rightarrow A^{*}A = \begin{bmatrix} 30 & 10 \\ 10 & 5 \end{bmatrix}$$

$$A^{*}\vec{b} = \begin{bmatrix} 32 \\ 12 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 30 & 10 \\ 10 & 5 \end{bmatrix} \vec{x} = \begin{bmatrix} 32 \\ 12 \end{bmatrix}$$

$$\Rightarrow \left[\begin{array}{cc|c} 30 & 10 & 32 \\ 10 & 5 & 12 \end{array}\right] \begin{array}{l} -3R2 \end{array} \Rightarrow \left[\begin{array}{cc|c} 0 & -5 & -4 \\ 10 & 5 & 12 \end{array}\right] \begin{array}{l} /-5 \\ /10 \end{array} \Rightarrow \left[\begin{array}{cc|c} 0 & 1 & 4/5 \\ 1 & 2 & 6/5 \end{array}\right]$$

$$\Rightarrow \hat{x} = \begin{bmatrix} 4/5 \\ 4/5 \end{bmatrix} \text{"="} \begin{bmatrix} a \\ b \end{bmatrix}.$$