Last class

Solving $Ax = b$ when A is tall

(and hence usually there is no solution)

$$J(x) = \| \overbrace{Ax - b} \|^2$$

$Ax - b$

residual

We try to make $J(x)$ as small as possible.

We are making the residual as small as possible.

If $\hat{x}$ is a minimizer of J

($J(\hat{x}) \leq J(x)$ for all other vectors x)

then $\hat{x}$ satisfies

$$Ax = b$$

$m \times n$ $\mathbb{R}^n$ $\mathbb{R}^m$

$m > n$

$$\underbrace{\overbrace{A^T A}^{\substack{n \times m \quad m \times n \\ n \times n}} \hat{x} = A^T b}_{\text{normal equation}}$$

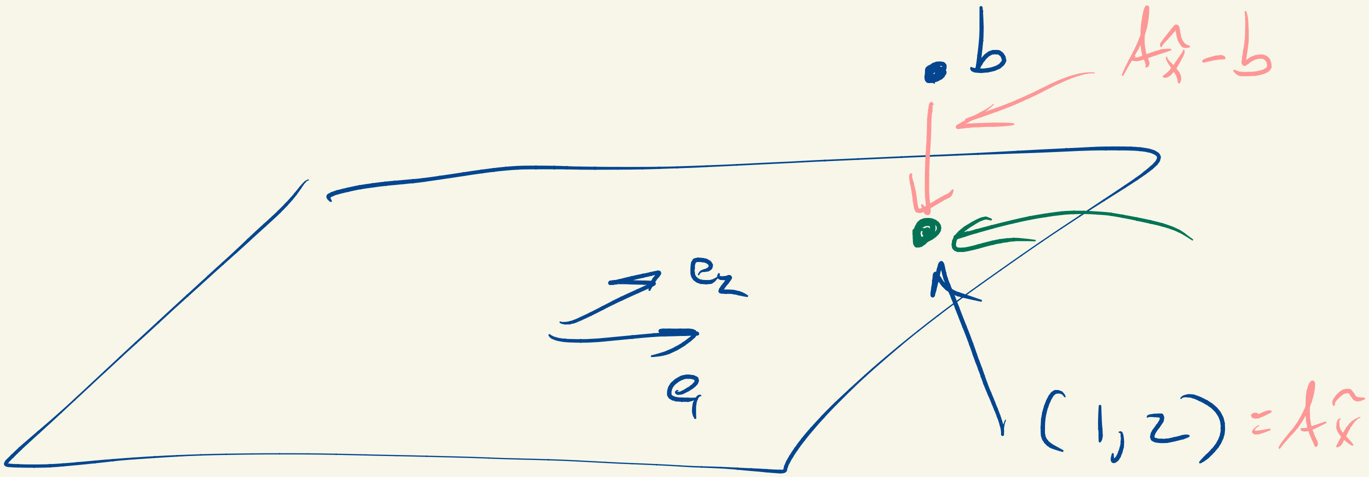
Assume columns of A are linearly independent

$A^T A$ is invertible.

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

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

$$\boxed{\hat{x} = A^{+} b}$$ (least squares solution)



$$\hat{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad Ax = b$$

$$A^T A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$A^T b$$

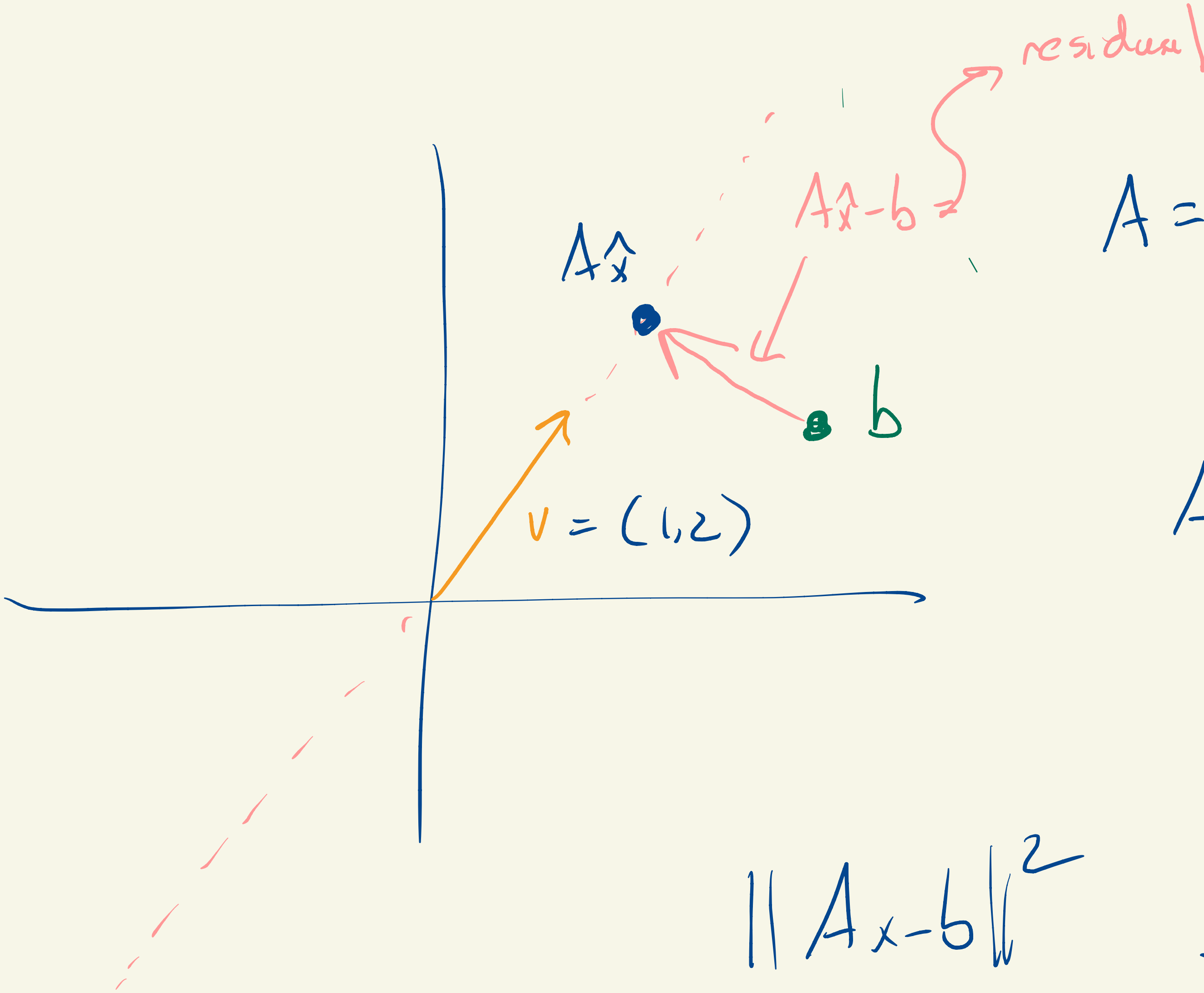
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

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

$$I \hat{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \Rightarrow \hat{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$\begin{bmatrix}2\\0\end{bmatrix} \to 4$



$A = \begin{bmatrix}1\\2\end{bmatrix}$

$J(1)=4$

$Ax = \begin{bmatrix}3\\2\end{bmatrix}$

↑
b

$\|Ax-b\|^2$

Task: find the point on the line as close as possible to b.

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

$$A = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad b = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$\underbrace{A^T A}_{5} \qquad \underbrace{A^T b}_{7}$$

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = 3 + 2 \cdot 2 = 7$$

$$5x = 7$$

$$x = 7/5$$

Closest point on line

$$\underbrace{\begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot \frac{7}{5} = \begin{bmatrix} 7/5 \\ 14/5 \end{bmatrix}}_{A\hat{x}}$$

Looks like we have the residual $A\hat{x}-b$

is orthogonal to v.

$$A\hat{x}-b = \begin{bmatrix} 7/5 \\ 14/5 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -8/5 \\ 4/5 \end{bmatrix}$$

↑ residual

$$v = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\frac{64}{25} + \frac{16}{25} = \frac{80}{25}$$

$$v^T(A\hat{x}-b) = -\frac{8}{5} + 2 \cdot \frac{4}{5} = 0 \checkmark$$

Claim: When you solve $A^T A \hat{x} = A^T b$

the residual $A\hat{x} - b$ is orthogonal

to any linear combination of the columns of A.

linear combo of columns of A: Az for some arbitrary vector z.

residual: $A\hat{x} - b$ $\qquad A^T A \hat{x} = A^T b$

$$(Az)^T (A\hat{x} - b) = z^T A^T (A\hat{x} - b)$$

$$= z^T \underbrace{\left(A^T A \hat{x} - A^T b\right)}_{0}$$

$= Z^T 0 = 0$

$Az \perp A\hat{x} - b$

How to solve

$$A^TAx = A^Tb$$

$$x = A^+ b$$

$$A = QR$$

1) $w = Q^Tb$

2) Solve $Rx = w$.

Then $x = A^+ b$

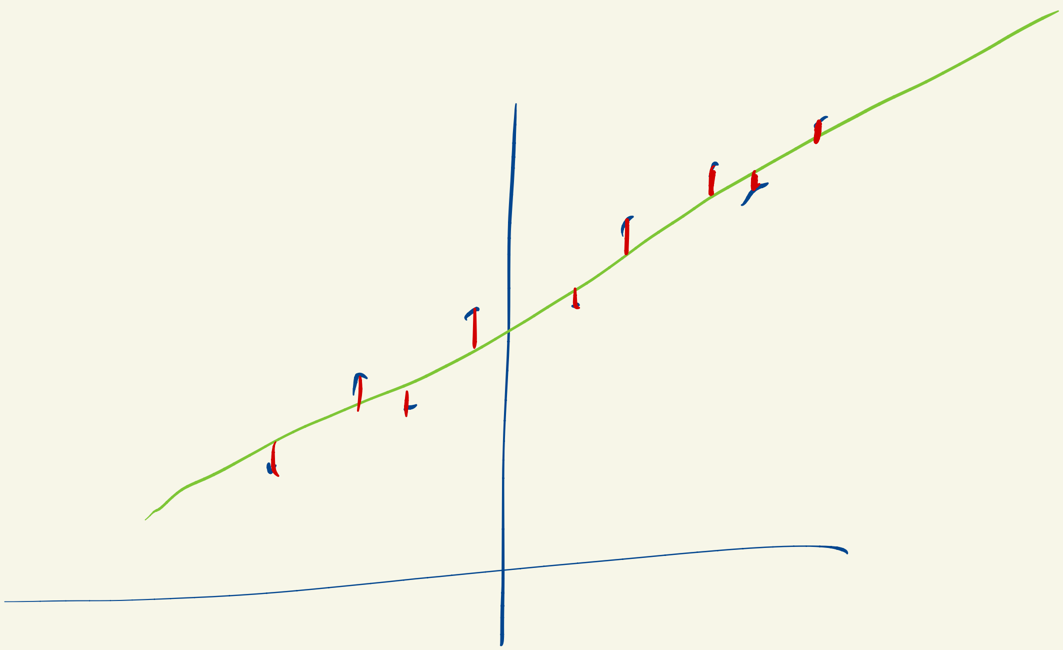
(x_1, y_1)

(x_2, y_2)

$\vdots$

(x_n, y_n)

Find m and b so that the line $y = mx + b$ passes through all these points.



$b + m x_1 = y_1$

$b + m x_2 = y_2$

$b + m x_3 = y_3$

$\vdots$

$b + m x_n = y_n$

$$\begin{bmatrix} 1 & x_1 \\ \vdots & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} b \\ m \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ \vdots \\ y_n \end{bmatrix}$$

We can look for a least squares solution

$$J((b,m)) = (y_1 - (b + m x_1))^2 + (y_2 - (b + m x_2))^2 + \cdots + (y_n - (b + m x_n))^2$$