$[\quad\quad] \leftarrow$ more than one solution.

Next class: How to compute A^+ using <u>QR</u> factorization for wide matrices.

$$A^T = QR \qquad\qquad A^+ = A^T(AA^T)^{-1}$$

$$AA^T = R^T \underbrace{Q^T Q}_{} R$$
$$= R^T R$$

$$(AA^T)^{-1} = (R^T R)^{-1}$$
$$= R^{-1}(R^T)^{-1}$$

$$A^+ = A^T R^{-1} (R^T)^{-1}$$

$$= Q \underbrace{R R^{-1}} (R^T)^{-1}$$

$$= Q (R^T)^{-1}$$

$$A^+ b = Q \overbrace{(R^T)^{-1} b}$$

1) Solve $\overbrace{R^T z = b}$

2) $x = Qz$

lower triangular

use forward subs.

$A^T = QR$

wide: $A^T (A A^T)^{-1}$

$z = (R^T)^{-1} b$

~~iff~~

$R^T z = b$

$A^T = QR$

upper

tri

Tall A: $\quad A^{+}b \quad\quad A = QR$

left inverse $\quad\quad A^{+} \; (A^TA)^{-1}A^T$

1) $z = Q^T b$

2) Solve $Rx = z$

What does $A^{+}b$ mean when A is tall and has linearly independent columns.

Least squares: $\quad Ax = b$

$\underbrace{A^{+}b}$ → "best value of x"

Ax is as close to b as possible.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Is there a solution?

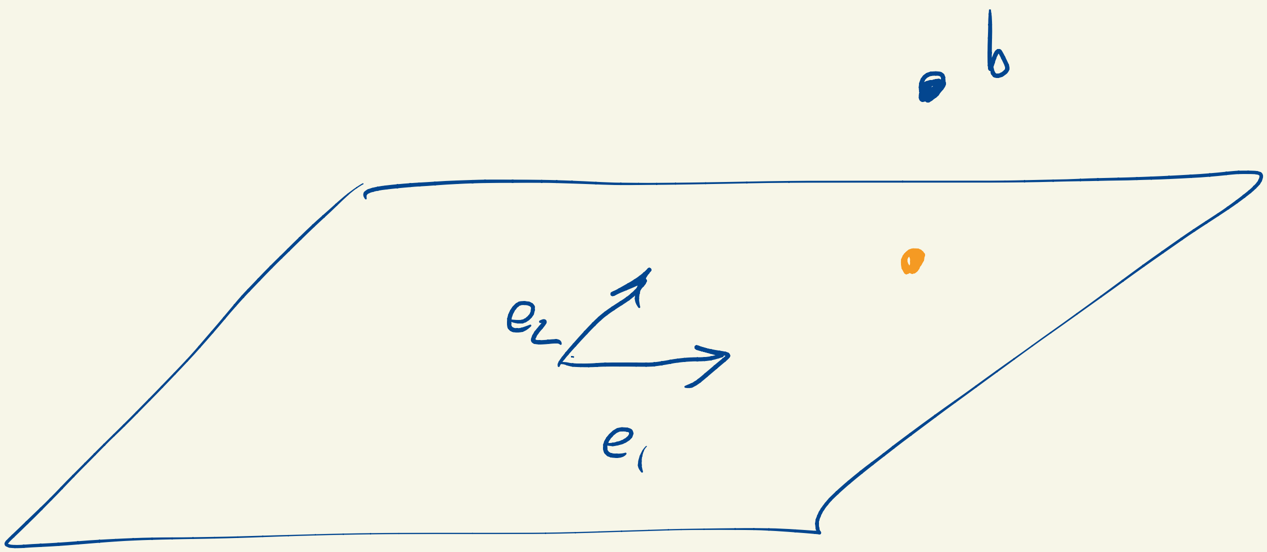
$$A x = b$$

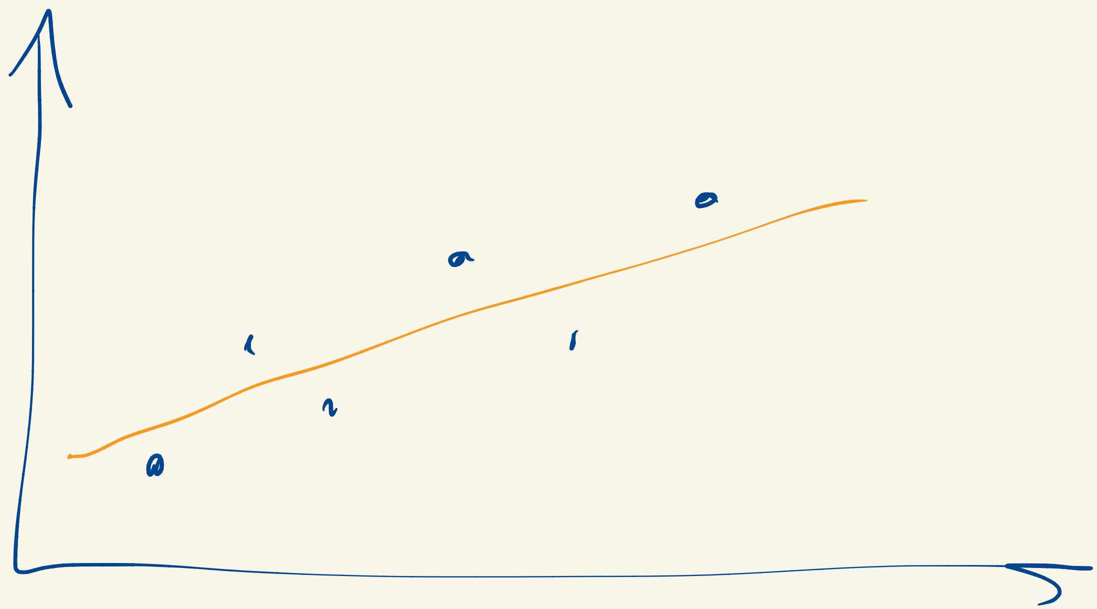
tall

lin. ind.

cols.

$$\begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$$





Task: Given the tall matrix A and the vector b, find $\hat{x}$ such that

$A\hat{x}$ is as close to b as possible.

(find the linear combination of the columns of A that is the best possible approximation of b.

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

"distance from Ax to b squared"

objective function

We want to find $\hat{x}$ such that $J(\hat{x})$ is as small as possible.

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

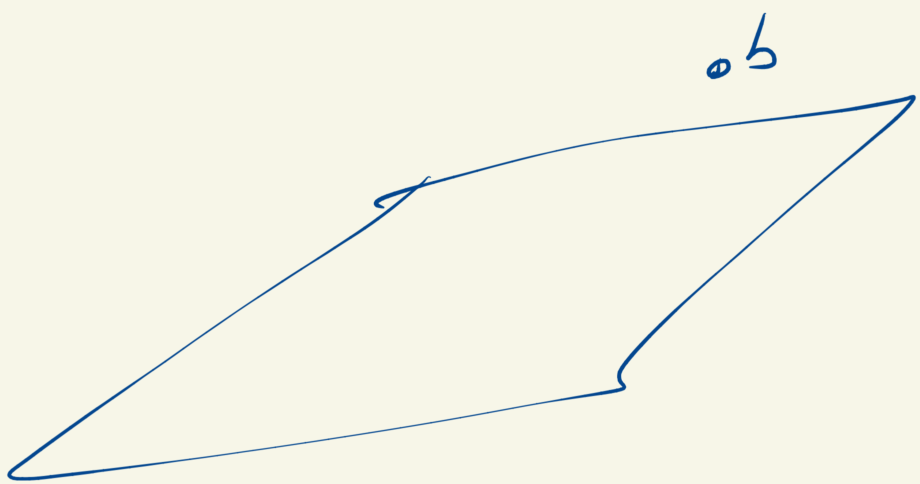
$$Ax = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} \quad Ax - b = \begin{bmatrix} x_1 - 1 \\ x_2 - 2 \\ -3 \end{bmatrix}$$

$$J(x) = \| Ax - b \|^2 = (x_1 - 1)^2 + (x_2 - 2)^2 + (-3)^2$$

$$= (x_1 - 1)^2 + (x_2 - 2)^2 + 9$$

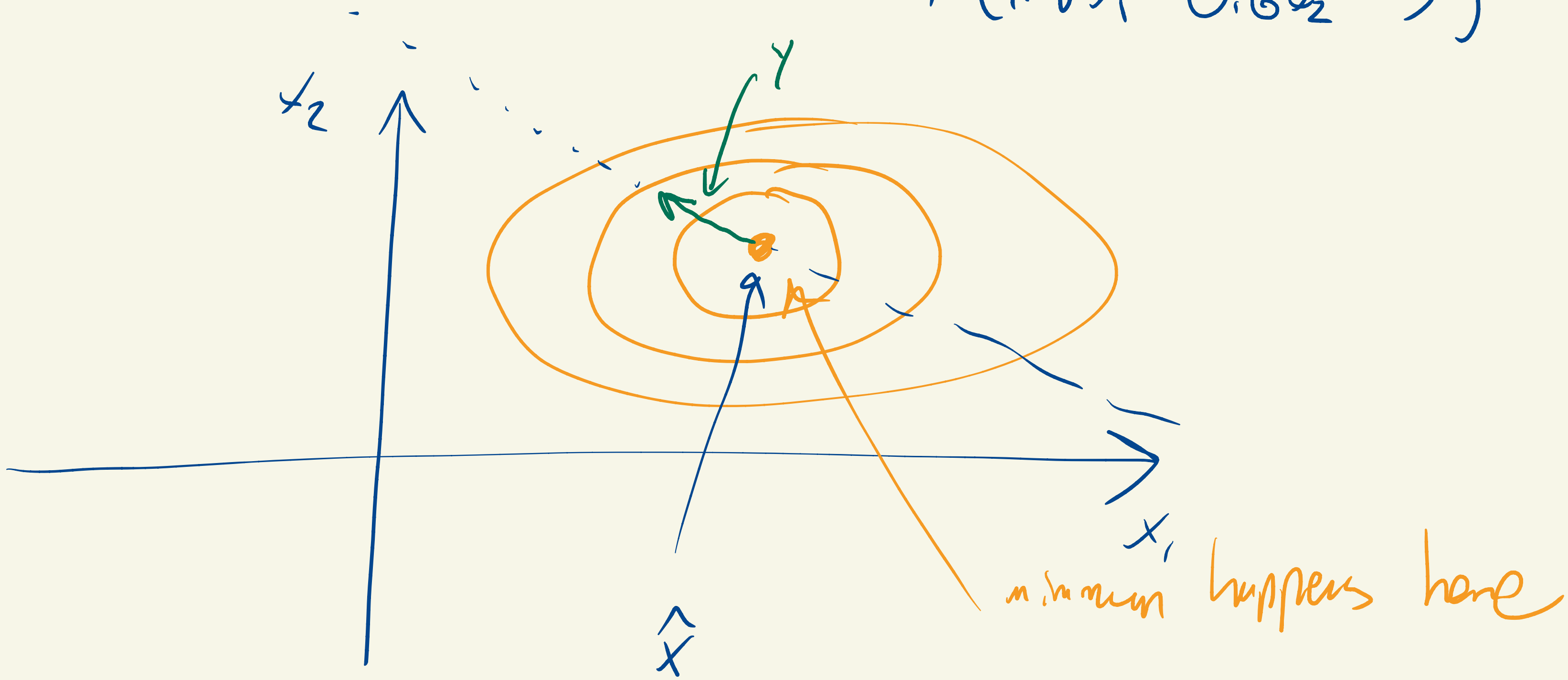
This is minimized when $x_1 = 1$ $x_2 = 2$.

$$A = \begin{bmatrix} 0.2 & 2.6 \\ -0.5 & 1.3 \\ 1.8 & -0.6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$



$Ax - b$ $\quad Ax - b = \begin{bmatrix} 0.2x_1 + 2.6x_2 - 1 \\ -0.5x_1 + 1.3x_2 - 2 \\ 1.8x_1 - 0.6x_2 - 3 \end{bmatrix}$

$$J(x) = \|Ax - b\|^2 = (0.2x_1 + 2.6x_2 - 1)^2 + (-0.5x_1 + 1.3x_2 - 2)^2 + (1.8x_1 - 0.6x_2 - 3)^2$$

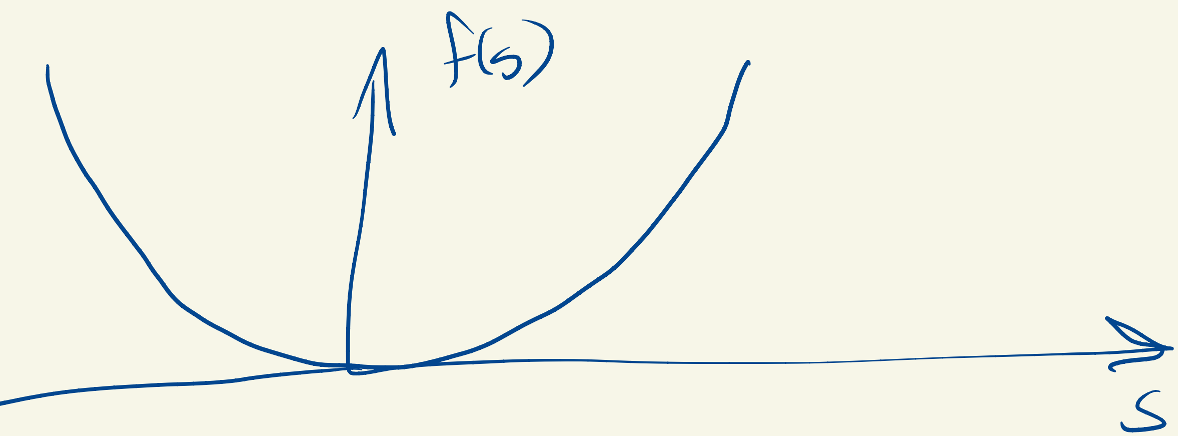


$\hat{x}$ is the location of the minimum,

y is some random direction

$\hat{x} + s y$

$$f(s) = J(\hat{x} + s y)$$



The minimum for f

happens when $f'(s) = 0$

(and the minimum is at $s=0$

$$f'(0) = 0$$

$$J(\hat{x} + s y) = \| A(\hat{x} + s y) - b \|^2$$

$$= \|(A\hat{x}-b)+sAy\|^2$$

$$\|v+w\|^2 = \|v\|^2+\|w\|^2+2v^Tw$$

$$\|v+w\|^2 = (v+w)^T(v+w)$$

$$= \|A\hat{x}-b\|^2 + 2s(Ay)^T(A\hat{x}-b) + s^2\|Ay\|^2$$

$$f(s)$$

$$f'(s) = 2(Ay)^T(A\hat{x}-b) + 2s\|Ay\|^2$$

$$f'(0) = 2(Ay)^T(A\hat{x}-b)$$

We need $2(\underline{Ay})^T(A\hat{x}-b) = 0$

for every direction y.

$$y^T \underbrace{A^T(A\hat{x}-b)} = 0 \quad \text{for all } y.$$

Suppose z is a vector and $y^T z = 0$ for all y.

$$\underbrace{z^T z} = 0$$

$$\|z\|^2 = 0 \qquad = 0$$

At a minimum $\hat{x}$ we have to have

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

$$\boxed{(A^TA)\hat{x} = A^Tb}$$

solve this

normal equation

We assume the columns of A are linearly independent.

Hence A^TA is invertible.

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

$$\underbrace{(A^TA)^{-1}A^T}$$

$\hookrightarrow A^{+}$ for the tall matrix A.

Finding $\hat{x}$ so that $A\hat{x}$ is as close a possible to b is exactly $\underbrace{\hat{x} = A^{+}b}$.

$$A = QR$$

1) $z = Q^T b$

2) $Rx = z$ by back subs.