6.5 Least-Squares Problems

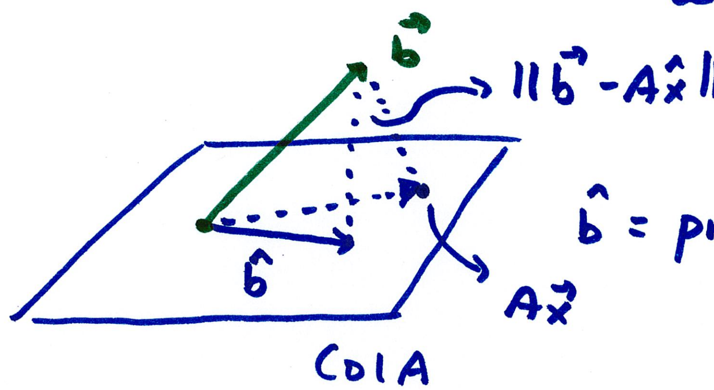
$$A = \begin{bmatrix} -1 & 4 \\ 2 & -3 \\ -1 & 3 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 12 \\ 0 \\ 6 \end{bmatrix}$$

$$\begin{bmatrix} A & \vec{b} \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \boxed{1} \end{bmatrix}$$

pivot in right most column $\rightarrow$ inconsistent

no solution to $A\vec{x} = \vec{b}$

$\vec{b}$ is not in Col A



$\hat{b} = \text{proj}_{\text{Col}A}\, \vec{b} = A\hat{x}$

some vector

so $A\hat{x} = \hat{b}$ <u>is</u> consistent

and $\|\vec{b} - A\hat{x}\| \le \|\vec{b} - A\vec{x}\|$

$\vec{x}$: some other $\vec{x}$

$A\vec{x}$: some vector in Col A

$\hat{x}$: Least-squares solution (magnitude is square root of sum of squares and minimized)

How to find $\hat{x}$?

$\vec{b} - \hat{b} = \vec{b} - A\hat{x}$ is orthogonal to Col A

so this means any column of A dotted with $\vec{b} - A\hat{x}$ is 0

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

$A^T\vec{b} - A^TA\hat{x} = \vec{0}$

$\boxed{A^TA\hat{x} = A^T\vec{b}}$ solve this for $\hat{x}$

$\vec{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ $\vec{c} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$

$\vec{a} \cdot \vec{c} = \vec{a}^T\vec{c}$

$B = \begin{bmatrix} 1 & -1 \\ 2 & -2 \\ 3 & -3 \end{bmatrix}$ $\vec{c} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$

$B^T\vec{c}$ is the same as dot products of columns of B w/ $\vec{c}$

<u>example</u> $A = \begin{bmatrix} -1 & 4 \\ 2 & -3 \\ -1 & 3 \end{bmatrix}$ $\vec{b} = \begin{bmatrix} 12 \\ 0 \\ 6 \end{bmatrix}$

$A^T = \begin{bmatrix} -1 & 2 & -1 \\ 4 & -3 & 3 \end{bmatrix}$

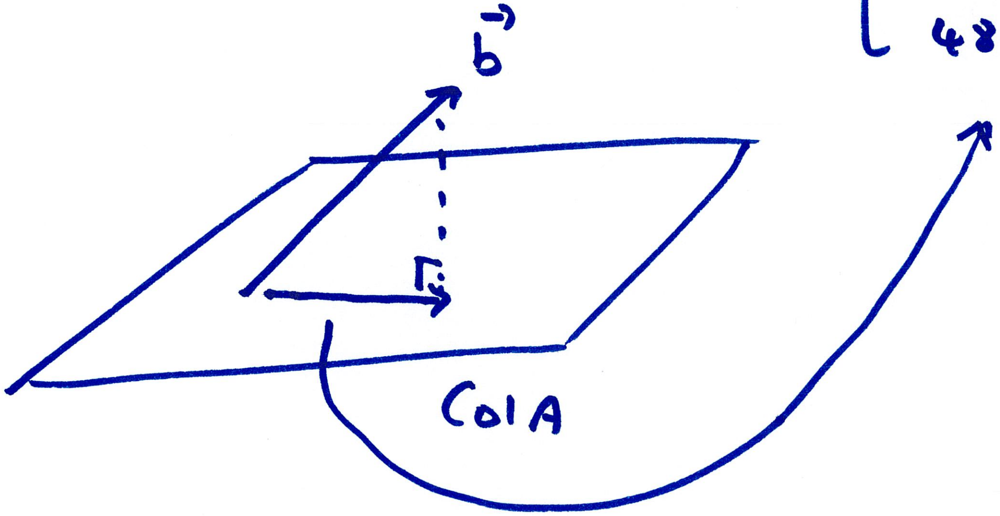
$A^TA = \begin{bmatrix} 6 & -13 \\ -13 & 34 \end{bmatrix}$ $A^T\vec{b} = \begin{bmatrix} -18 \\ 66 \end{bmatrix}$

$\begin{bmatrix} 6 & -13 & -18 \\ -13 & 34 & 66 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 246/35 \\ 0 & 1 & 162/35 \end{bmatrix}$

$$\hat{x} = \begin{bmatrix} 246/35 \\ 162/35 \end{bmatrix}$$

and $$\hat{b} = \text{proj}_{\text{Col}A}\vec{b} = \frac{246}{35}\begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix} + \frac{162}{35}\begin{bmatrix} 4 \\ -3 \\ 3 \end{bmatrix}$$

$$= A\hat{x} = \begin{bmatrix} 402/35 \\ 6/35 \\ 48/7 \end{bmatrix}$$



$A^T A \hat{x} = A^T \vec{b}$ **ALWAYS** has at least a solution

(because ~~A~~ $\hat{b}$ is in Col A)

but $\hat{x}$ can have multiple solutions example

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 2 \\ 2 \\ 3 \\ 7 \end{bmatrix}$$

$$[A \;\; \vec{b}\,] \sim \cdots \sim \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

pivot in right most column

$A\vec{x} = \vec{b}$ has no solution

$\vec{b}$ is not in Col A

find $\hat{\vec{x}}$: $A^TA\,\hat{x} = A^T\vec{b}$

$\hookrightarrow$ also works if $A\vec{x} = \vec{b}$ is consistent

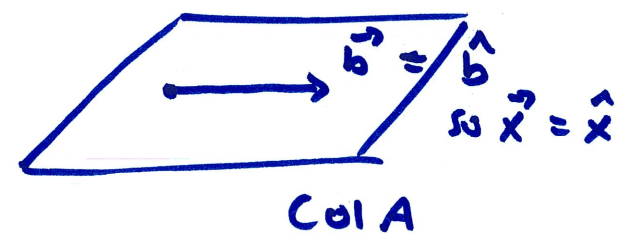
$$A^TA = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix}$$

A: $n \times m$

A^T: $m \times n$

A^TA = $m \times m$

always square



$$A^T\vec{b} = \begin{bmatrix} 14 \\ 4 \\ 10 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 2 & 2 & 14 \\ 2 & 2 & 0 & 4 \\ 2 & 0 & 2 & 10 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 1 & 5 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

x_3 free

$x_2 = x_3 - 3$

$x_1 = 5 - x_3$

$$\hat{x} = \begin{bmatrix} 5 \\ -3 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

there are many ways to build $\hat{b}$ out of columns of

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

because the columns of A here are not linearly independent

$$A^TA = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix}$$ has the same structure

so A^TA is not invertible

so $(A^TA)\hat{x} = A^T\vec{b}$ does not have unique solution

if columns of A $\underline{\underline{\text{are}}}$ linearly independent

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

problematic if A is large or messy

because small errors in A^TA can make $\hat{x}$ very wrong to stabilize the algorithm, use $A = QR$

R → upper triangular matrix w/ pos. diagonal elements

Q → columns are orthonormal basis for Col A

$$Q^T Q = I \rightarrow Q^T = Q^{-1}$$

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

$$(QR)^T QR \hat{x} = (QR)^T \vec{b}$$

$$R^T \underbrace{Q^T Q}_{I} R \hat{x} = R^T Q^T \vec{b}$$

$$R^T R \hat{x} = R^T Q^T \vec{b}$$

$$R \hat{x} = \underbrace{(R^T)^{-1} R^T}_{I} Q^T \vec{b}$$

R is triangular w/ nonzero main diagonal, so $(R^T)^{-1}$ exists.

$$\boxed{\hat{x} = R^{-1} Q^T \vec{b}}$$