

Approximate Linear equations

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

Sometimes no solns

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

$\vec{b} \notin \text{col}(A)$, i.e.

\vec{b} is not a LC of

$$\begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}.$$

How close can we get?

make $A\vec{x}$ as close to \vec{b} as possible.

form us, this means minimize

$$\|A\vec{x} - \vec{b}\| = \sqrt{(A\vec{x} - \vec{b}) \cdot (A\vec{x} - \vec{b})}$$

$$= \sqrt{(Ax_1 - b_1)^2 + (Ax_2 - b_2)^2 + \dots + (Ax_n - b_n)^2}$$

Other option: minimize

$$(|Ax_1 - b_1| + |Ax_2 - b_2| + \dots + |Ax_n - b_n|)$$

L₁ norm

Dfn: $A \in M_{m \times n}$ matrix, $\vec{b} \in \mathbb{R}^m$.

A least-squares soln to $A\vec{x} = \vec{b}$

\exists a vector $\hat{\vec{x}} \in \mathbb{R}^n$ s.t.

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

$\forall \vec{x} \in \mathbb{R}^n$.

Can min w/ calc
but that's gross.

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

want to minimize

$$\left\| x_1 \begin{bmatrix} 4 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 4x_1 - 2 \\ 2x_2 \\ x_1 + x_2 - 1 \end{bmatrix} \right\| = \sqrt{(4x_1 - 2)^2 + (2x_2)^2 + (x_1 + x_2 - 1)^2},$$

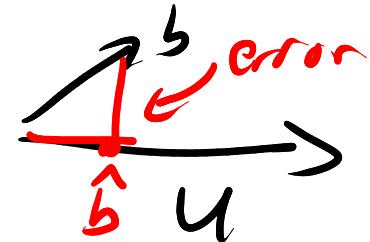
$A\vec{x}$

The set $\left\{ x_1 \begin{bmatrix} 4 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 2 \end{bmatrix} \right\} = U$

is a 2-d subspace.

Want distance from $\vec{b} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ to U .

project \vec{b} onto U .



Orth basis for U : $\left\{ \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ -2/5 \end{bmatrix} \right\}$

$$\text{proj}_{\vec{u}_1} \vec{b} = \frac{\vec{b} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 = \frac{11}{5} \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

$$\text{proj}_{\vec{u}_2} \vec{b} = \frac{84/5}{84/5} \begin{bmatrix} 4 \\ -2/5 \end{bmatrix} = \begin{bmatrix} 4 \\ -2/5 \end{bmatrix}$$

$$\begin{aligned} \text{proj}_{\text{def}} \vec{b} &= \frac{11}{5} \begin{bmatrix} 0 \\ 2 \end{bmatrix} + \frac{1}{5} \begin{bmatrix} 20 \\ -2 \end{bmatrix} \\ &= \frac{1}{5} \begin{bmatrix} 20 \\ 15 \end{bmatrix} - \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \hat{b}. \end{aligned}$$

so $\begin{bmatrix} 4 \\ 0 \\ 3 \end{bmatrix}$ is the closest I can get to $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$.

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

$$\hat{\vec{x}} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 4 & 0 & 4 \\ 0 & 2 & 0 \\ 1 & 1 & 3 \end{array} \right] \rightarrow \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 3 \end{array} \right]$$

Idea was: project \vec{b} onto $\text{col}(A)$ to get $\hat{\vec{b}}$

instead of solving $A\vec{x} = \vec{b}$

solve $A\vec{x} = \hat{\vec{b}}$, $\hat{\vec{b}} \in U = \text{col}(A)$.

Suppose have a soln

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

$\vec{b} - \hat{\vec{b}}$ is the error

and is \perp to U .

so $\vec{b} - \hat{\vec{b}}$ is \perp to
every column of A .

↓

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

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

$$\text{so } A^T(\vec{b}) = A^T \hat{\vec{b}}$$

$$A = \begin{bmatrix} \vec{c}_1 & \dots & \vec{c}_n \end{bmatrix}$$

$$A^T = \begin{bmatrix} \vec{c}_1 \\ \vdots \\ \vec{c}_n \end{bmatrix}$$

$$A^T \vec{y} = \begin{bmatrix} \vec{c}_1 \cdot \vec{y} \\ \vec{c}_2 \cdot \vec{y} \\ \vdots \\ \vec{c}_n \cdot \vec{y} \end{bmatrix}$$

Prop: the set of least-squares solns to $A\vec{x} = \vec{b}$

is the set of solns to $\underbrace{A^T A \vec{x}}_{\text{normal equations}} = \underbrace{A^T \vec{b}}_{\text{right side}}$.

Ex: $\begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \vec{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

Solve $\begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 \end{bmatrix} \vec{x} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \vec{x} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

$$\begin{bmatrix} 17 & 1 & | & 19 \\ 1 & 5 & | & 11 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & | & 1 \\ 0 & 1 & | & 2 \end{bmatrix}$$

Pf:
If $A\vec{x} = \vec{b}$ then
 $A^T A\vec{x} = A^T \vec{b} = \vec{b}$ (last slide)

Conversely, suppose \vec{x} solves $A^T A\vec{x} = \vec{b}$
then $\underline{A^T(\vec{b} - A\vec{x})} = 0$

so $\vec{b} - A\vec{x}$ is \perp to $\text{col } A$. But

$\vec{b} = \underset{\text{col } A}{\vec{x}} + \underset{\text{col } A^\perp}{(\vec{b} - A\vec{x})}$ is ortho decomp.
thus $\vec{x} = \underset{\text{col } A}{\text{proj}_{\text{col } A} \vec{b}}$.

So, least squares
soln is $\vec{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Ex: find a least-squares soln to

$$\begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \\ -1 & -1 \end{bmatrix} \vec{x} = \begin{bmatrix} 2 \\ 5 \\ 6 \\ 6 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

$$A^T \vec{b} = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 6 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ 14 \\ -5 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 14 \\ -5 \end{bmatrix}$$

$$\text{So } \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1/3 \\ 14/3 \\ -5/3 \end{bmatrix}.$$

$$\text{Moreover, } \vec{b} = A \vec{x} = \begin{bmatrix} 5 \\ 2 \\ 3 \\ 6 \end{bmatrix}$$

$$\vec{b} - \vec{b} = \begin{bmatrix} -3 \\ 3 \\ 3 \end{bmatrix}$$

Magnitude of error is $\|\vec{b} - \vec{b}\| = \sqrt{22}$.

\hat{x} is unique when $A^T A$ is invertible

iff cols of A are LI.

For this case, $\hat{x} = (A^T A)^{-1} A^T \vec{b}$

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

doesn't work b/c A not square!

Regression.

have input data X
output data \bar{y} .

want 'parameter vector' β that
makes $X\beta \approx \bar{y}$ as close as
possible

x_1, y_1
 x_2, y_2
;
 x_n, y_n

want to find
 $\beta_0 + \beta_1 x \approx y$.

Set $X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$, $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$, $\bar{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$

want to find least-squares sol.
to $X\bar{\beta} = \bar{y}$.

$(2, 3)$
 $(3, 2)$
 $(5, 1)$
 $(6, 0)$

$$X = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 5 \\ 1 & 6 \end{bmatrix} \quad \bar{y} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \end{bmatrix}$$

LS soln to $X\vec{\beta} = \vec{y}$

Model:
 $y = \beta_0 + \beta_1 x$

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 5 \\ 1 & 6 \end{bmatrix} = \begin{bmatrix} 4 & 16 \\ 16 & 74 \end{bmatrix}$$

$$X^T \bar{y} = \begin{bmatrix} 6 \\ 17 \end{bmatrix}$$

line of best fit is

$$y = 4.3 - .7x$$

$$\left[\begin{array}{cc|c} 4 & 16 & 6 \\ 16 & 74 & 17 \end{array} \right] \rightarrow \left[\begin{array}{cc|c} 1 & 0 & 4.3 \\ 0 & 1 & -.7 \end{array} \right].$$

model blood pressure as a function of weight by $P = \beta_0 + \beta_1 \ln(w)$

$$\begin{matrix} P & 91 & 98 & 103 & 110 & 112 \\ w & 44 & 61 & 81 & 113 & 131 \end{matrix}$$

$$\ln w \quad 3.78 \quad 4.11 \quad 4.39 \quad 4.73 \quad 4.88$$

$$X = \begin{bmatrix} 1 & 3.78 \\ 1 & 4.11 \\ 1 & 4.39 \\ 1 & 4.73 \\ 1 & 4.88 \end{bmatrix} \quad \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 91 \\ 98 \\ 103 \\ 110 \\ 112 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 5.00 & 21.89 \\ 21.89 & 96.64 \end{bmatrix}$$

$$X^T \bar{y} = \begin{bmatrix} 514 \\ 2265.79 \end{bmatrix}$$

$$\begin{bmatrix} 5.00 & 21.89 & 514 \\ 21.89 & 96.64 & 2265.79 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 18.57 \\ 0 & 1 & 19.24 \end{bmatrix}$$

$$P = 18.57 + 19.24 \ln(w)$$

What, if I want best cubic?

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$$

$$\begin{array}{l} (2,3) \\ (3,2) \\ (5,1) \\ (6,0) \end{array} \quad \left[\begin{array}{cccc} 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 5 & 25 & 125 \\ 1 & 6 & 36 & 216 \\ 12 & 29 & 7^3 \\ 110 & 100 & 10^{100} \end{array} \right] \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \\ -2 \\ -4 \end{bmatrix}$$