Math 225 lecture 16th oct 13th 2023

Section 7.3

Goal: students should be able to compute a LSS via two different methods. Moreover they should be able to identify when to employ a LSS as opposed to solving a system normaly.

Class Q: when approaching a problem (system in this case) when do you need to use a least squares solution?

Recall: last time we showed that if presented with a system $A\vec{x}=\vec{b}$ which cannot be solved there is a notion of "best fit" solution found by solving $A\vec{x}=\vec{p}$ w/ $\vec{p}=\text{proj}_{\text{col}(A)}(\vec{b})$ (noting that we need an orthogonal basis for col(A)) is there a better way?

Yes! set up as before $A\vec{x}=\vec{b}$ not solvable.

replace with $A\vec{x}=\vec{p}$ and solve

(recall: $\vec{b}=\vec{p}+\vec{p}^{\perp}$)

$$\Rightarrow \vec{b}-\vec{p}=\vec{p}^{\perp}$$
$$\vec{b}-A\vec{x}=\vec{p}^{\perp}\in(\text{Col}(A))^{\perp}$$

(from lecture 9 sept 25th $\text{Col}(A)^{\perp}=\text{null}(A^{T})$)

there for since $(\vec{b}-A\vec{x})\in\text{Col}(A)^{T}$ we have

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

$$\Rightarrow \boxed{A^{T}\vec{b}=A^{T}A\vec{x}}$$

this is called the "normal system of best fit problem" or "least squares fit" things to note:

- If A's columns are linearly independant then A^TA will be invertible & $\therefore$ $\vec{x} = (A^TA)^{-1} \cdot A^T\vec{b}$ has a unique sol'n

- (A^TA) is invertible but each of A & A^T are <u>not</u>

- This defines another method of solving for projections, namely
  $\text{proj}_{col(A)}(\vec{b}) = \underbrace{A \cdot (A^TA)^{-1} A^T \vec{b}}$ (~~this is done as proj is the least squares soln~~
  note this is just $A\vec{x}$

- this proposed projection works independant of $\vec{b}$

Ex: $\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$ (this is the one we did last class as we saw the answer should be $\begin{bmatrix} -1 \\ 5/2 \end{bmatrix}$

lets do this via the normal system, this requires 2 pieces $(A^TA)^{-1}$ & $A^T\vec{b}$

$$(A^TA)^{-1} = \left(\begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \right)^{-1} = \left(\begin{bmatrix} 6 & 4 \\ 4 & 3 \end{bmatrix} \right)^{-1} = \frac{1}{2} \begin{bmatrix} 3 & -4 \\ -4 & 6 \end{bmatrix}$$

$$= \begin{bmatrix} 3/2 & -2 \\ -2 & 3 \end{bmatrix}$$

$$A^T\vec{b} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 11 \\ 7 \end{bmatrix}$$

$\vec{x} = (A^TA)^{-1} A^T\vec{b} = \begin{bmatrix} 3/2 & -2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5/2 \\ -1 \end{bmatrix}$ slight path

Who remembers the QR factorization?

<u>Recall</u>: We factor $A = QR$ with Q ? orthogonal and R ? upper triangular

- apply G-S to the columns of A & normalize these form the columns of Q

- $R = Q^TA$

- we have the factorization $A = QR$ ✓

this also plays a role here!

<u>claim</u>: $\vec{x} = R^{-1}Q^T\vec{b}$ where $\vec{x}$ is our LSS

<u>proof</u>: $A^TA\vec{x} = A^T\vec{b}$ also $A = QR$ use this here

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

$R^TQ^TQR\vec{x} = R^TQ^T\vec{b}$ since Q is orthogonal

$R^TR\vec{x} = R^TQ^T\vec{b}$

since R is upper triangular & invertible R^T will also be invertible (why?)

$\Rightarrow R\vec{x} = Q^T\vec{b}$ or $\vec{x} = R^{-1}Q^T\vec{b}$

$\boxtimes$

Ex: the QR method.

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

if you apply G-S to the columns of A you get

$$Q = \begin{bmatrix} 1/2 & \frac{3\sqrt{5}}{10} & -\sqrt{6}/6 \\ -1/2 & \frac{3\sqrt{5}}{10} & 0 \\ -1/2 & \sqrt{5}/10 & \sqrt{6}/6 \\ 1/2 & \sqrt{5}/10 & \sqrt{6}/3 \end{bmatrix}$$

multiplying $Q^T A$ we see that

$$R = \begin{bmatrix} 2 & 1 & 1/2 \\ 0 & \sqrt{5} & 3\sqrt{5}/2 \\ 0 & 0 & \sqrt{6}/2 \end{bmatrix}$$

at this point we can invert R & use $\vec{x} = R^{-1} Q^T \vec{b}$ or solve $R\vec{x} = Q^T \vec{b}$

$$Q^T \vec{b} = \begin{bmatrix} 1/2 & -1/2 & -1/2 & 1/2 \\ \frac{3\sqrt{5}}{10} & 3\sqrt{5}/10 & \sqrt{5}/10 & \sqrt{5}/10 \\ -\sqrt{6}/6 & 0 & \sqrt{6}/6 & \sqrt{6}/3 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 7/2 \\ -\sqrt{5}/2 \\ -2\sqrt{6}/3 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 & 1/2 \\ 0 & \sqrt{5} & 3\sqrt{5}/2 \\ 0 & 0 & \sqrt{6}/2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7/2 \\ -\sqrt{5}/2 \\ -2\sqrt{6}/3 \end{bmatrix}$$

Hilroy

at this point we can quickly back substitute and solve (this is likely easier than inverting R)

$$\begin{aligned} 2x_1 + x_2 + \tfrac{1}{2}x_3 &= \tfrac{7}{2} \\ \sqrt{5}x_2 + \tfrac{3\sqrt{5}}{2}x_3 &= -\tfrac{\sqrt{5}}{2} \\ \tfrac{\sqrt{6}}{2}x_3 &= \tfrac{-2\sqrt{6}}{3} \end{aligned}$$

$$\Rightarrow \vec{x} = \begin{bmatrix} 4/3 \\ 3/2 \\ -4/3 \end{bmatrix}$$