Linear Algebra Recitation

Week 7

Jun Chang

New York University

October 22, 2021

Best approximation (Least sqares)

- Let $A = \begin{bmatrix} 1 & -1 \\ -1 & 2 \\ -1 & 0 \end{bmatrix}$ and $b = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$ Does $Ax = b$ have a solution $x \in \mathbb{R}^2$?

$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$$\underbrace{\begin{bmatrix} 1 & -1 \\ -1 & 2 \\ -1 & 0 \end{bmatrix}}_{A} \Bigg| \underbrace{\begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}}_{b} \sim \left[\begin{array}{cc|c} 1 & -1 & 3 \\ 0 & 1 & 7 \\ 0 & -1 & 8 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -1 & 3 \\ 0 & 1 & 7 \\ \underline{0} & \underline{0} & \underline{15} \end{array}\right]$$

$$\Rightarrow x_1 - x_2 = 3$$

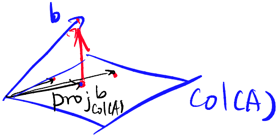
$$\left.\begin{array}{l} x_2 = 7 \\ x_2 = -8 \end{array}\right) (\lightning)$$

Best approximation (Least sqares)

- For a system $Ax = b$, a least squares solution, or the best approximation, gives the vector in col(A) that is closest to b (i.e. minimizes $\|b - Ax\|_2$). Thus, in geometric terms, the least squares solution denoted by $\hat{x}$ is the coordinates in terms of the basis of col(A) that gives the projection vector of b, namely $A\hat{x} = \text{Proj}_{\text{col}(A)} b$.

$Ax = \text{proj}_{\text{Col}(A)} b = \hat{b}$

$Ax = b$



$b - A\hat{x} \in \text{Col}(A)^{\perp}$

- How to find $\hat{x}$: If $\hat{x}$ is the least squares solution, the residual $r = b - A\hat{x}$ must be orthogonal to every vector in col(A). In particular, the following must hold:

$$\begin{bmatrix} c_1 \cdot (b - A\hat{x}) \\ \vdots \\ c_n \cdot (b - A\hat{x}) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$A^T = \begin{bmatrix} c_1^T \\ \vdots \\ c_n^T \end{bmatrix}$$

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

- Thus, we can solve the new system of equations $A^T A\hat{x} = A^T b$ to find $\hat{x}$.

Best approximation (Least sqares) $Ax=b$

- Does $A^T A\hat{x} = A^T b$ (a.k.a. normal equation) have a solution? Notice that the right hand side is $A^T b \in \text{col}(A^T)$
- If $\text{col}(A^T) = \text{col}(A^T A) \implies$ there will be at least one solution!

  Indeed, $A^TAv = A^Tw$ $\quad$ $\text{rank}(A^T) = \text{rank}(A^TA)$

  $\text{Col}(A^TA) \subset \text{Col}(A^T)$ $\quad$ $\Rightarrow \text{Col}(A^TA) = \text{Col}(A^T)$

- If $A^T A$ has full rank, it is invertible and the least squares solution is

$\text{Proj}_{\text{Col}(A)} b = A\hat{x}$

$= A(A^TA)^{-1}A^Tb$

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

Unique

- What if $A^T A$ is not full rank?: The solution $\hat{x}$ is not unique. How do we find a least squares solution $\hat{x}$ in this case?

  $\implies$ not a part of this course, but we use the concept of "pseudoinverses".

Best approximation (Least sqares)

Back to the example problem:

- Let $A = \begin{bmatrix} 1 & -1 \\ -1 & 2 \\ -1 & 0 \end{bmatrix}$ and $b = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$ Find the least squares solution: namely, find $\hat{x}$ that minimizes $\|b - Ax\|_2$.

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

$A^T = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 2 & 0 \end{bmatrix}$

$A^TA = \begin{bmatrix} 3 & -3 \\ -3 & 5 \end{bmatrix}$

$A^Tb = \begin{bmatrix} -6 \\ 5 \end{bmatrix}$

$\begin{bmatrix} 3 & -3 \\ -3 & 5 \end{bmatrix} x = \begin{bmatrix} -6 \\ 5 \end{bmatrix}$

$\left[\begin{array}{cc|c} 3 & -3 & -6 \\ 0 & 2 & -1 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -1 & -2 \\ 0 & 1 & -1/2 \end{array}\right]$

$x_1 - x_2 = -2$

$x_2 = -1/2$

$x_1 = -5/2$

$\hat{x} = \begin{bmatrix} -5/2 \\ -1/2 \end{bmatrix}$

Uniqueness of least squares solution

- The least squares solution is unique when $A^T Ax = A^T b$ has a unique solution $\iff$ $\mathcal{N}(A^TA) = \{\vec{0}\}$ (trivial)
- When is $\mathcal{N}(A^T A)$ trivial? (How is it related to $\mathcal{N}(A)$)?

① $v \in N(A^TA) \Rightarrow A^TAv = 0 \Rightarrow v^TA^TAv = 0 \Rightarrow \|Av\|^2 = 0 \Rightarrow Av = 0$
$\Rightarrow v \in N(A) \Rightarrow N(A^TA) \subset N(A)$

② $u \in N(A) \Rightarrow Au = 0 \Rightarrow A^TAu = 0 \Rightarrow u \in N(A^TA) \Rightarrow N(A) \subset N(A^TA)$

- What happens if the normal equation has infinitely many solutions? In particular, does the value Ax change among the solutions x of the normal equation?

x, y be least squares solutions $A^TAx = A^Tb = A^TAy$
is $Ax = Ay$? $x^* = x + x_N$, $x_N \in N(A^TA) = N(A)$
$Ax^* = Ax + Ax_N = Ax$ so $\text{proj}_{Col(A)} b$ is unique
even though x^* is not unique.

$$\underbrace{\begin{bmatrix} 1 & 0 \\ -1 & 0 \\ -1 & 0 \end{bmatrix}}_{A} x = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}}_{b}$$

$$\underbrace{\begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}}_{A^T A} x = \underbrace{\begin{bmatrix} -4 \\ 0 \end{bmatrix}}_{A^T b}$$

$3x_1 = -4 \Rightarrow x_1 = -4/3$

$0x_2 = 0$

$x_N \in N(A^T A)$

$$x = x_p + x_N$$

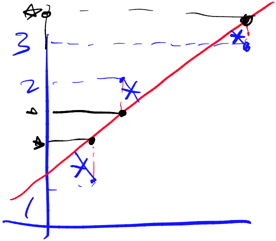
$$= \begin{bmatrix} -4/3 \\ 0 \end{bmatrix} + x_N$$

infinitely many least squares solutions.

Application of least squares:

Fitting a line through data.

- Suppose we want to find the line that best fits (in terms of euclidean distance) three data points (1,1), (2,2), (4,3) in $\mathbb{R}^2$. How to find this line using least squares?
- *If* there is a line $\overset{f(x)}{y} = ax + b$ that interpolates these points exactly, what would this line do?



$$f(1) = a + b = 1$$
$$f(2) = 2a + b = 2$$
$$f(4) = 4a + b = 3$$

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Application of least squares:

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 21 & 7 \\ 7 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 17 \\ 6 \end{bmatrix}, \quad \left[\begin{array}{cc|c} 7 & 3 & 6 \\ 21 & 7 & 17 \end{array}\right] \sim \left[\begin{array}{cc|c} 7 & 3 & 6 \\ 0 & -2 & -1 \end{array}\right]$$

$$7a + 3b = 6 \qquad b = 1/2$$

$$a = \frac{1}{7}\left(9/2\right) = 9/14$$

$$y = \frac{9}{14}x + 1/2$$