Lecture 35. Least squares problems

<u>Def</u> Given an equation $A\vec{x}=\vec{b}$ for an $m\times n$ matrix A and a vector $\vec{b}\in\mathbb{R}^m$, its <u>least squares solution</u> is a vector $\hat{x}\in\mathbb{R}^n$ with

$$\|A\hat{x}-\vec{b}\|\le\|A\vec{x}-\vec{b}\| \text{ for any } \vec{x}\in\mathbb{R}^n.$$

<u>Note</u> (1) If the equation $A\vec{x}=\vec{b}$ is solvable, $\hat{x}$ is an actual solution of the equation.

(2) If the equation $A\vec{x}=\vec{b}$ is not solvable, $\hat{x}$ is an approximate solution of the equation with minimum error $\|A\hat{x}-\vec{b}\|$.

<u>Thm</u> Given an equation $A\vec{x}=\vec{b}$ for a matrix A and a vector $\vec{b}$, its least squares solution $\hat{x}$ is given by the equation $A^TA\hat{x}=A^T\vec{b}$.

<u>pf</u> $\mathrm{Col}(A)$ is the set of all vectors of the form $A\vec{x}$.

$\hat{x}$ is a least squares solution

$\iff \|A\hat{x}-\vec{b}\|\le\|A\vec{x}-\vec{b}\|$ for any $\vec{x}$

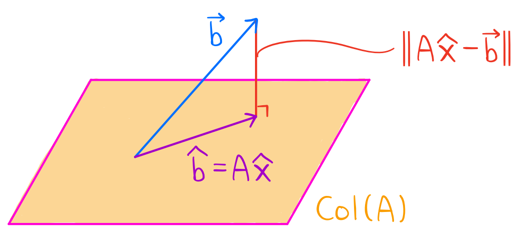
$\iff A\hat{x}$ is the closest vector to $\vec{b}$ in $\mathrm{Col}(A)$

$\iff A\hat{x}$ is the orthogonal projection of $\vec{b}$ onto $\mathrm{Col}(A)$

$\iff A\hat{x}-\vec{b}$ lies in $\mathrm{Col}(A)^\perp=\mathrm{Nul}(A^T)$

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

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



Ex Find the least squares solution of the linear system

$$\begin{cases} x_1 + x_2 = 1 \\ x_1 - x_2 = 2 \\ x_1 + 3x_2 = 6 \end{cases}$$

Sol The linear system can be written as $A\vec{x} = \vec{b}$ with

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 3 \end{bmatrix} \text{ and } \vec{b} = \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix}.$$

The least squares solution $\hat{x}$ is given by the equation $A^T A \hat{x} = A^T \vec{b}$.

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 11 \end{bmatrix}$$

$$A^T \vec{b} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix} = \begin{bmatrix} 9 \\ 17 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 3 & 3 & 9 \\ 3 & 11 & 17 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 1 \end{array}\right]$$

$A^T A \quad A^T \vec{b}$

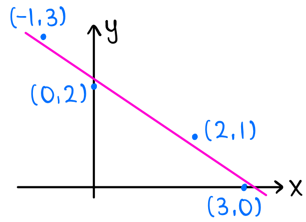
$$\Rightarrow \boxed{x_1 = 2, \ x_2 = 1}$$

Note Alternatively, we may compute

$$\hat{x} = (A^T A)^{-1} A^T \vec{b} = \frac{1}{3 \cdot 11 - 3 \cdot 3} \begin{bmatrix} 11 & -3 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} 9 \\ 17 \end{bmatrix} = \frac{1}{24} \begin{bmatrix} 48 \\ 24 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

<u>Ex</u> Find the equation of the least squares line that best fits the data given in the table

x	-1	0	2	3
y	3	2	1	0



<u>Sol</u> We write $y = \alpha + \beta x$ for the least squares line.

We want an equation $A\vec{x} = \vec{b}$ for $\vec{x} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$.

$$\begin{cases} 3 = \alpha - \beta & (x=-1,\ y=3) \\ 2 = \alpha & (x=0,\ y=2) \\ 1 = \alpha + 2\beta & (x=2,\ y=1) \\ 0 = \alpha + 3\beta & (x=3,\ y=0) \end{cases} \Longrightarrow A = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \text{ and } \vec{b} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \end{bmatrix}.$$

The least squares solution $\hat{x}$ is given by the equation $A^T A \hat{x} = A^T \vec{b}$.

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 4 & 14 \end{bmatrix}$$

$$A^T \vec{b} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 2 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ -1 \end{bmatrix}$$

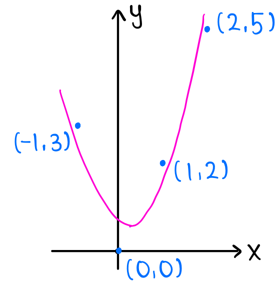
$$\left[\begin{array}{cc:c} 4 & 4 & 6 \\ 4 & 14 & -1 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{cc:c} 1 & 0 & 2.2 \\ 0 & 1 & -0.7 \end{array}\right]$$

$A^T A \quad A^T \vec{b}$

$$\Longrightarrow \alpha = 2.2,\ \beta = -0.7 \Longrightarrow \boxed{y = 2.2 - 0.7x}$$

Ex Find the least squares quadratic function $f(x)$ that best fits the data given in the table

x	-1	0	1	2
y	3	0	2	5



Sol We write $f(x) = \alpha + \beta x + \gamma x^2$.

We want an equation $A\vec{x} = \vec{b}$ for $\vec{x} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}$.

$$\begin{cases} 3 = \alpha - \beta + \gamma & (x=-1,\ y=3) \\ 0 = \alpha & (x=0,\ y=0) \\ 2 = \alpha + \beta + \gamma & (x=1,\ y=2) \\ 5 = \alpha + 2\beta + 4\gamma & (x=2,\ y=5) \end{cases} \Rightarrow A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \text{ and } \vec{b} = \begin{bmatrix} 3 \\ 0 \\ 2 \\ 5 \end{bmatrix}.$$

The least squares solution $\hat{x}$ is given by the equation $A^T A \hat{x} = A^T \vec{b}$.

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix}$$

$$A^T \vec{b} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 10 \\ 9 \\ 25 \end{bmatrix}$$

$$\underbrace{\left[\begin{array}{ccc|c} 4 & 2 & 6 & 10 \\ 2 & 6 & 8 & 9 \\ 6 & 8 & 18 & 25 \end{array}\right]}_{A^T A \quad A^T\vec{b}} \xrightarrow{\text{RREF}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0.6 \\ 0 & 1 & 0 & -0.7 \\ 0 & 0 & 1 & 1.5 \end{array}\right]$$

$$\Rightarrow \alpha = 0.6,\ \beta = -0.7,\ \gamma = 1.5 \Rightarrow f(x) = \boxed{0.6 - 0.7x + 1.5x^2}$$