భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad
॥ नवोन्मेषेण देश निर्माणम् ॥

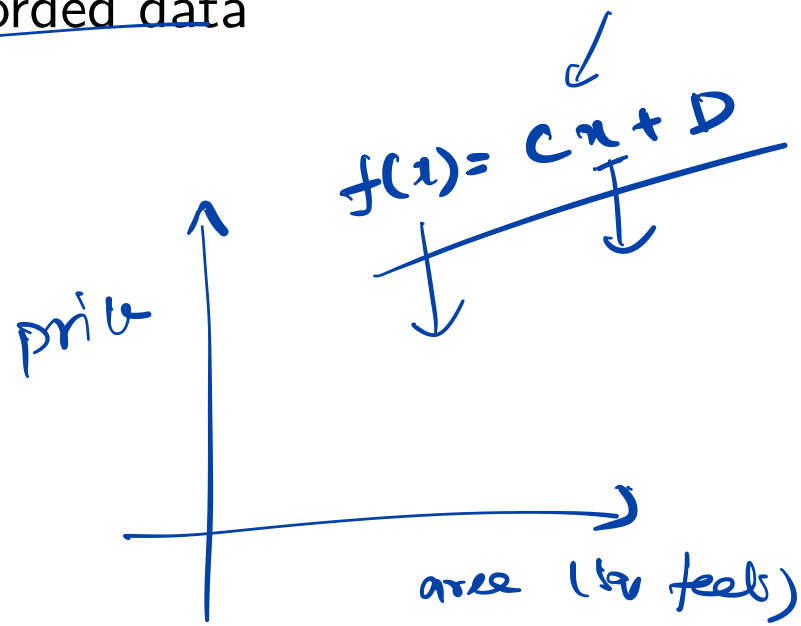
Matrix Theory

16 LS Fitting of Data

Dr. Konda Reddy Mopuri
Dept. of AI, IIT Hyderabad
July-Nov 2026

LS Fitting of Data

- Via a series of m experiments, we recorded data



LS Fitting of Data

- Via a series of m experiments, we recorded data
- Expectation is that the output b is a linear function of i/p t



LS Fitting of Data

$$\left.\begin{aligned} C + Dt_1 &= b_1 \\ C + Dt_2 &= b_2 \\ &\cdots\cdots \\ C + Dt_m &= b_m \end{aligned}\right\}$$

- How to compute C and D?

LS Fitting of Data

$$C + Dt_1 = b_1$$
$$C + Dt_2 = b_2$$
$$\cdots\cdots$$
$$C + Dt_m = b_m$$

- How to compute C and D?
- Over-determined system (m equations with 2 variables)

LS Fitting of Data

$$C + Dt_1 = b_1$$
$$C + Dt_2 = b_2$$
$$\cdots\cdots$$
$$C + Dt_m = b_m$$

- How to compute C and D?
- Over-determined system (m equations with 2 variables)
- If there are errors, there won't be a solution

LS Fitting of Data

target

$$\begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \dots & \dots \\ 1 & t_m \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}$$

A x b

- Best solution $(\hat{C}, \hat{D})$ is the one that minimizes the SSE
  $E^2 = \|Ax - b\|^2 = \sum_i^m (C + Dt_i - b_i)^2$

LS Fitting of Data

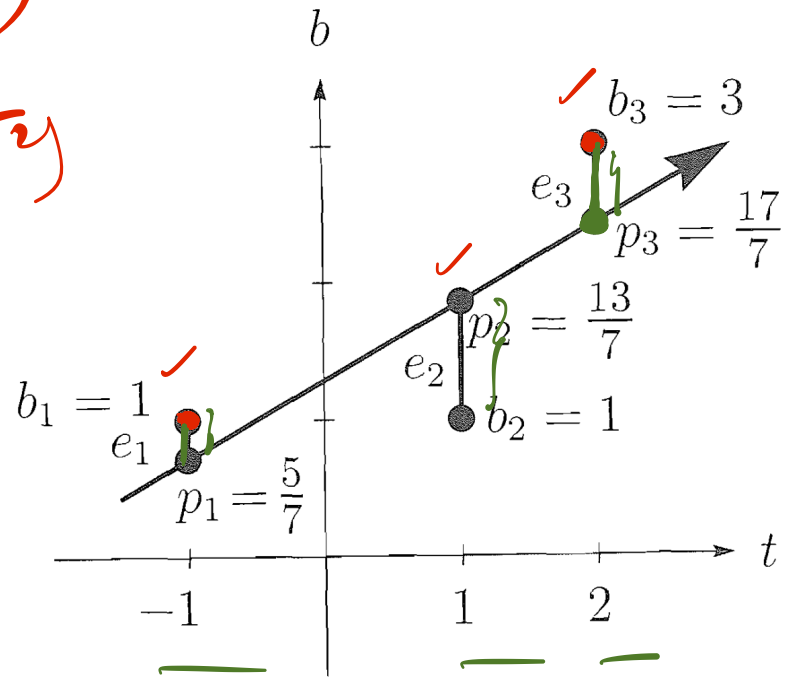
$\hat{C}$ & $\hat{D}$

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

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$$

$A \quad x \quad b$

LS Fitting of Data



Straight-line Approximation and the error terms (Note: not orthogonal to the fitted line)

LS Fitting of Data

- These equations can't be solved

LS Fitting of Data

- These equations can't be solved
- Therefore the LS solution: $A^T A\hat{x} = A^T b$

LS Fitting of Data

$$\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$$

- Best solution is $\hat{C} = \frac{9}{7}$ and $\hat{D} = \frac{4}{7}$

LS Fitting of Data

$$\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$$

- Best solution is $\hat{C} = \frac{9}{7}$ and $\hat{D} = \frac{4}{7}$
- Best line is $\hat{b} = \frac{9}{7} + \frac{4}{7}t$

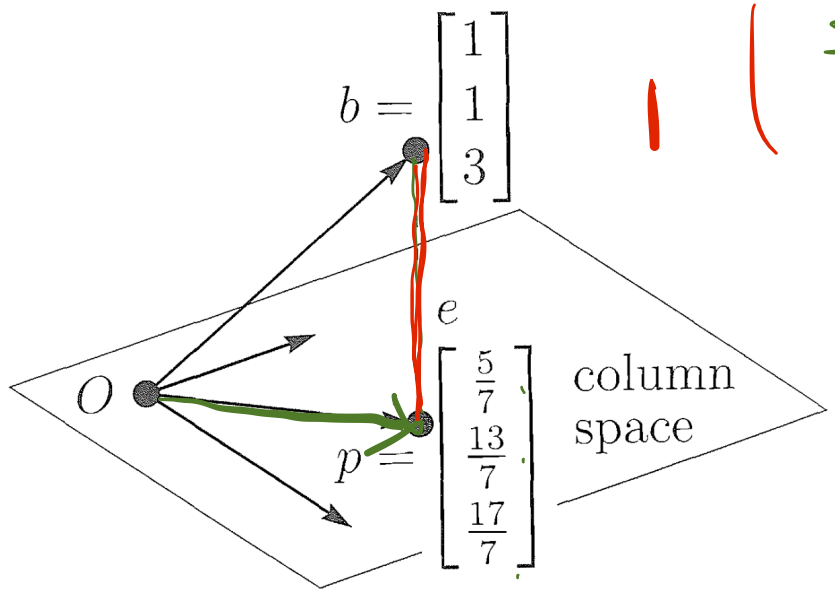
$A\hat{x} = \hat{b}$

$\hat{b} = C + Dt + Et^2 + Ft^3$

LS Fitting of Data

$$\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$$

- Best solution is $\hat{C} = \frac{9}{7}$ and $\hat{D} = \frac{4}{7}$
- Best line is $\hat{b} = \frac{9}{7} + \frac{4}{7}t$



Projection of b onto the column space of A

Weighted Least Squares

- What if the observations/measurements are of different importance?

Weighted Least Squares

- What if the observations/measurements are of different importance?
- Some of them are obtained in 'more' reliable way than others

Weighted Least Squares

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} x = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

- Until now, we treated b_1 and b_2 equally reliable

Weighted Least Squares

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} x = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

- Until now, we treated b_1 and b_2 equally reliable
- Looked for $\hat{x}$ that minimizes the SSE: $E^2 = (x - b_1)^2 + (x - b_2)^2$

Weighted Least Squares

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} x = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

- Until now, we treated b_1 and b_2 equally reliable
- Looked for $\hat{x}$ that minimizes the SSE: $E^2 = (x - b_1)^2 + (x - b_2)^2$
- Optimal $\hat{x}$ is the average of the b_1 and b_2

Weighted Least Squares

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} x = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

- Let's attach different weights w_1^2 and w_2^2 to the errors

Weighted Least Squares

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} x = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

- Let's attach different weights w_1^2 and w_2^2 to the errors
- Weighted error $E^2 = w_1^2(x - b_1)^2 + w_2^2(x - b_2)^2$

$w_1 > w_2$

Weighted Least Squares

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} x = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

- Let's attach different weights w_1^2 and w_2^2 to the errors
- Weighted error $E^2 = w_1^2(x - b_1)^2 + w_2^2(x - b_2)^2$
- If $w_1 > w_2$, the optimization process will try to make the first term small

Weighted Least Squares

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} x = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

- Let's attach different weights w_1^2 and w_2^2 to the errors
- Weighted error $E^2 = w_1^2(x - b_1)^2 + w_2^2(x - b_2)^2$
- If $w_1 > w_2$, the optimization process will try to make the first term small
- $\frac{dE^2}{dx} = 2[w_1^2(x - b_1) + w_2^2(x - b_2)]$ and $\hat{x}_w = \frac{w_1^2 b_1 + w_2^2 b_2}{w_1^2 + w_2^2}$

Weighted Least Squares

LS normal equation $A^T Ax = A^T b$

$\|Ax-b\|^2$

- LS solution becomes:

Weighted Least Squares

- LS solution becomes:
- Weighted Normal equations become:

$$(WA)^T (WA) x = (WA)^T b$$

Weighted Least Squares

$r = (A\hat{x} - b)$

- LS solution becomes:
- Weighted Normal equations become:
- Note that the notion of distance and orthogonality changes

$A\hat{x}_w$ ⊥?

$(A\hat{x}_w - b)$ or Something else

Rough Work

$$\|A\hat{x} - b\|^2$$

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

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

$$\hat{x}^T A^T A \hat{x} - 2\hat{x}^T A^T b + b^T b$$

$$2A^T A\hat{x} - 2A^T b = 0$$

Oct- 2, Fri
11 - 1:10 pm
Midsem

post Q2 till LS

50 - 60 %.

Rest before Q2