Alternative derivation

$$E = \sum_{j=1}^{n} \left(m x_j + c - y_j \right)^2$$

---

idea: write αE as $(...)^2 + \underbrace{(...)(...)}_{\oplus}^2 + (...)$

$\min(\alpha E) \longrightarrow$ obtained by setting to zero

---

notation: $Z = \sum_{j=1}^{n} y_j^2$

$$\alpha E = \alpha \sum_{j=1}^{n} \left(m x_j + c - y_j \right)^2$$

$$= \alpha \sum_{j=1}^{n} \left(m^2 x_j^2 + c^2 + y_j^2 - 2c y_j + 2mc x_j - 2m x_j y_j \right)$$

$$= \alpha m^2 + \alpha n c^2 + \alpha Z - 2\alpha c \delta + 2mc \alpha \beta - 2m\alpha \gamma$$

$$= \alpha^2 m^2 + 2\alpha m \left(\beta c - \gamma \right) + \alpha \left(n c^2 + Z - 2\delta c \right)$$

$$\underbrace{}_{\substack{\text{complete} \ldots \\ \text{in } \alpha m}}$$

$$= \left(\alpha m + \beta c - \gamma \right)^2 - \underbrace{\left(\beta c - \gamma \right)^2}_{\beta^2 c^2 + \gamma^2 - 2\beta \gamma c} + \alpha \left(n c^2 + Z - 2\delta c \right)$$

$$= \left(\alpha m + \beta c - \gamma \right)^2 + \left(\alpha n - \beta^2 \right) c^2 + 2 \left(\beta \gamma - \alpha \delta \right) c - \gamma^2 + \alpha Z$$

$$= \left(\alpha m + \beta c - \gamma \right)^2 + \left(\alpha n - \beta^2 \right) \left[c^2 + 2 \frac{\beta \gamma - \alpha \delta}{\alpha n - \beta^2} \right] - \gamma^2 + \alpha Z$$

$$= (\alpha m + \beta c - \gamma)^2 + (\alpha n - \beta^2)\left[c + \frac{\beta\gamma - \alpha\delta}{\alpha n - \beta^2}\right]^2 + \alpha Z - \frac{(\beta\gamma - \alpha\delta)^2}{\alpha n - \beta^2} - \gamma^2$$

$\oplus$

$$\alpha n - \beta^2 = n\left(\sum x_j^2\right) - \left(\sum x_j\right)^2 \geq 0 \quad [\text{always}]$$

αE minimum $\rightarrow$
$$\begin{cases} \alpha m + \beta c - \gamma = 0 \,^{\circledast} & \leftarrow \text{---} \; 1^{st} \; \text{normal equation} \\ c + \dfrac{\beta\gamma - \alpha\delta}{\alpha n - \beta^2} = 0 \end{cases}$$

$$(\alpha n - \beta^2)c + \beta\gamma - \alpha\delta = 0$$
$$\alpha n c - \alpha\delta - \beta^2 c + \beta\gamma = 0$$
$$\alpha n c - \alpha\delta = \underbrace{\beta^2 c - \beta\gamma}_{\beta(\beta c - \gamma)\,\overset{\circledast}{=}\, - \beta\alpha m}$$

$$\alpha n c - \alpha\delta = -\beta\alpha m \; \Big) : \alpha \;\Rightarrow\; \boxed{\beta m + n c = \delta} \;\; 2^{nd} \; \text{normal equation}$$

$$\boxed{E_{min} = Z - \frac{1}{\alpha}\left[\frac{(\beta\gamma - \alpha\delta)^2}{\alpha n - \beta^2} + \gamma^2\right]}$$

Matrix formulation for least squares fit

Normal equations → $\begin{bmatrix} \alpha & \beta \\ \beta & n \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} \gamma \\ \delta \end{bmatrix}$

earlier notation:

$$\alpha = \sum_{j=1}^{n} x_j^2 \qquad \beta = \sum_{j=1}^{n} x_j$$

$$\gamma = \sum_{j=1}^{n} x_j y_j \qquad \delta = \sum_{j=1}^{n} y_j$$

Define:

$$A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \qquad B = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \qquad Z = \begin{bmatrix} m \\ c \end{bmatrix}$$

Vector of unknowns

$$A^T = \begin{bmatrix} x_1 & x_2 \cdots & 1 \\ 1 & 1 \cdots & 1 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} \alpha & \beta \\ \beta & n \end{bmatrix} \qquad A^T B = \begin{bmatrix} \gamma \\ \delta \end{bmatrix}$$

$$\boxed{A^T A Z = A^T B}$$

Example

x	$\overset{x_1}{0}$	$\overset{x_2}{1}$	$\overset{x_3}{1}$	$\overset{x_4}{2}$	$\overset{x_5}{2}$
y	0	0	1	2	3
	y_1	y_2	y_3	y_4	y_5

what is the best straight line that fits this data in the least squares sense?

Solution

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 1 \\ 2 & 1 \\ 2 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 0 & 1 & 1 & 2 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 1 \\ 2 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 6 \\ 6 & 5 \end{bmatrix}$$
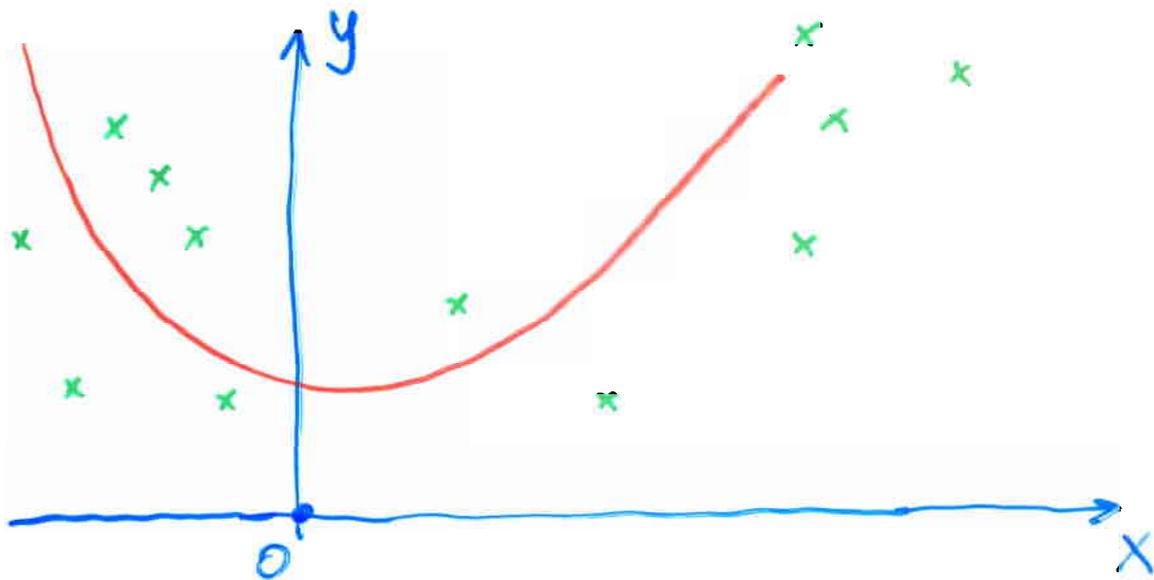
$$A^T B = \begin{bmatrix} 0 & 1 & 1 & 2 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 11 \\ 6 \end{bmatrix}$$

normal eqns. $\implies \begin{bmatrix} 10 & 6 \\ 6 & 5 \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} 11 \\ 6 \end{bmatrix} \implies$

$$M = \frac{19}{4}$$
$$c = -\frac{6}{14}$$

$\implies$ best fit is the line $\boxed{y = \frac{19}{14} x - \frac{3}{7}}$
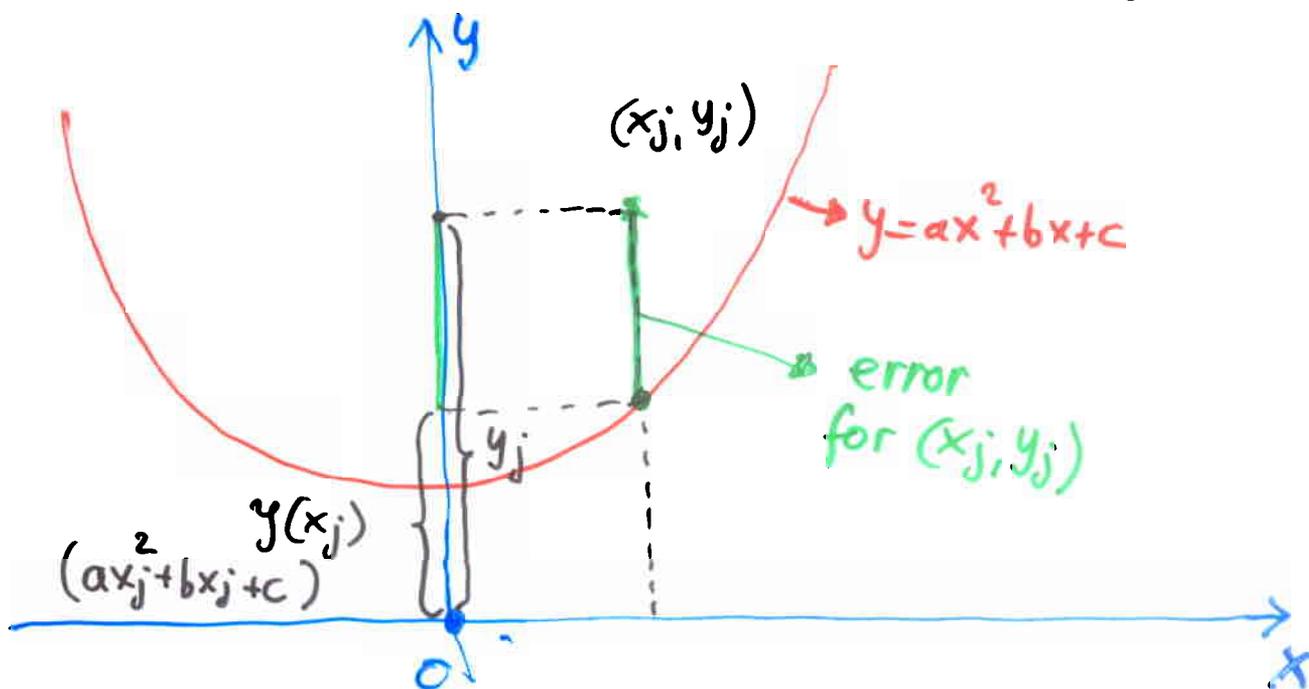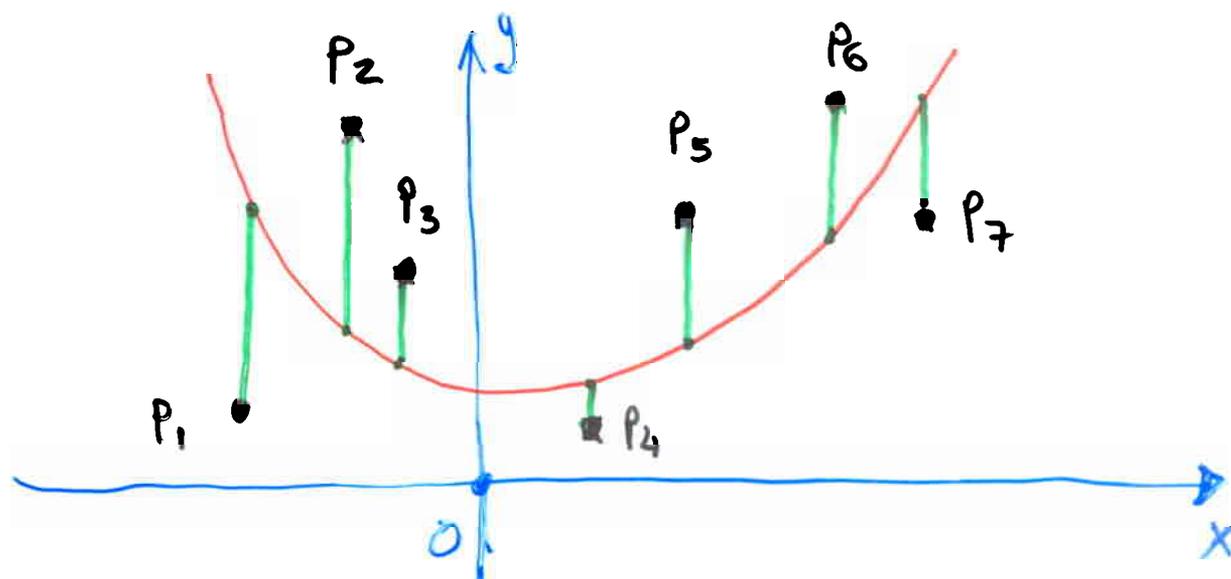
Fitting a parabola

Main question:

Find the "best" parabola

$$y = ax^2 + bx + c$$

which fits a given set of points

OBS. every parabola is uniquely determined by 3 numbers
(a, b, c)

(x_j, y_j)

$y = ax^2 + bx + c$

error
for (x_j, y_j)

y_j

$y(x_j)$
$(ax_j^2 + bx_j + c)$

$$P_j \equiv (x_j, y_j) \qquad 1 \leq j \leq n \qquad \text{set of data points}$$

$$E_j = |y(x_j) - y_j| = |a x_j^2 + b x_j + c - y_j|$$

"best fit" $\longrightarrow$ $\boxed{\displaystyle\sum_{j=1}^{n} E_j^2 \text{ as small as possible}}$ $\quad \circledast$
(least squares)

$$\underbrace{\sum_{j=1}^{n} E_j^2}_{E}$$

$$E = \sum_{j=1}^{n} \left(a x_j^2 + b x_j + c - y_j \right)^2 \longrightarrow \text{depends on 3 arbitrary parameters } a, b, c$$

$$E = E(a, b, c)$$

$\circledast \Rightarrow \boxed{\text{Find } a, b, c \text{ s.t. } E(a, b, c) \text{ is minimum}}$

$E \geq 0 \longrightarrow$ possible critical points will be minima of E

$$\frac{\partial E}{\partial a}(a,b,c) = 0 \qquad \frac{\partial E}{\partial b}(a,b,c) = 0 \qquad \frac{\partial E}{\partial c}(a,b,c) = 0$$

find a, b, c s.t. these equations are simultaneously satisfied

$$\frac{\partial E}{\partial a} = \sum_{j=1}^{n} \frac{\partial}{\partial a} \left\{ (ax_j^2 + bx_j + c - y_j)^2 \right\}$$

$$= \sum_{j=1}^{n} 2(ax_j^2 + bx_j + c - y_j) \underbrace{\frac{\partial}{\partial a} \left\{ (ax_j^2 + bx_j + c - y_j) \right\}}_{x_j^2}$$

$$= \sum_{j=1}^{n} 2(ax_j^2 + bx_j + c - y_j) x_j^2$$

$$\frac{\partial E}{\partial b} = \sum_{j=1}^{n} \frac{\partial}{\partial b} \left\{ (ax_j^2 + bx_j + c - y_j)^2 \right\}$$

$$= \sum_{j=1}^{n} 2(ax_j^2 + bx_j + c - y_j) x_j$$

$$\frac{\partial E}{\partial c} = \sum_{j=1}^{n} \frac{\partial}{\partial c} \left\{ (ax_j^2 + bx_j + c - y_j)^2 \right\}$$

$$= \sum_{j=1}^{n} 2(ax_j^2 + bx_j + c - y_j) \cdot 1$$

OBS. We are after a, b, c (and not x_j or y_j)

3 linear equations in 3 unknowns

$$\begin{cases} \sum_{j=1}^{n} \left(a x_j^2 + b x_j + c - y_j \right) x_j^2 = 0 \\ \sum_{j=1}^{n} \left(a x_j^2 + b x_j + c - y_j \right) x_j = 0 \\ \sum_{j=1}^{n} \left(a x_j^2 + b x_j + c - y_j \right) = 0 \end{cases}$$

$\Rightarrow$

$$\begin{cases} a \underbrace{\left(\sum x_j^4 \right)}_{\beta_1} + b \underbrace{\left(\sum x_j^3 \right)}_{\beta_2} + c \underbrace{\left(\sum x_j^2 \right)}_{\beta_3} = \underbrace{\sum x_j^2 y_j}_{\delta_1} \\ a \underbrace{\left(\sum x_j^3 \right)}_{\beta_2} + b \underbrace{\left(\sum x_j^2 \right)}_{\beta_3} + c \underbrace{\left(\sum x_j \right)}_{\beta_4} = \underbrace{\sum x_j y_j}_{\delta_2} \\ a \underbrace{\left(\sum x_j^2 \right)}_{\beta_3} + b \underbrace{\left(\sum x_j \right)}_{\beta_4} + c \underbrace{n}_{\beta_5} = \underbrace{\sum y_j}_{\delta_3} \end{cases}$$

$$\Downarrow$$

$$\begin{pmatrix} \beta_1 & \beta_2 & \beta_3 \\ \beta_2 & \beta_3 & \beta_4 \\ \beta_3 & \beta_4 & \beta_5 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{pmatrix} \Rightarrow \text{use Gaussian elimination, etc}$$

$\beta_1, \beta_2, \dots, \beta_5 \longrightarrow$ known
from the data

hard to remember $\longrightarrow$ $\longrightarrow$ $\longrightarrow$

$$A = \begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & & \\ x_n^2 & x_n & 1 \end{bmatrix} \underbrace{\qquad}_{n \times 3} \rightarrow A^T = \begin{bmatrix} x_1^2 & x_2^2 & \cdots & x_n^2 \\ x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \underbrace{\qquad}_{3 \times n}$$

$$A^T A \rightarrow 3 \times 3$$

$$\left[A A^T \rightarrow n \times n \right]$$

$$A^T A = \begin{bmatrix} \sum_{j=1}^{n} x_j^4 & \sum_{j=1}^{n} x_j^3 & \sum_{j=1}^{n} x_j^2 \\ \sum_{j=1}^{n} x_j^3 & \sum_{j=1}^{n} x_j^2 & \sum_{j=1}^{n} x_j \\ \sum_{j=1}^{n} x_j^2 & \sum_{j=1}^{n} x_j & \underbrace{\sum 1}_{n} \end{bmatrix}$$

Let $B = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \longrightarrow A^T B = \begin{bmatrix} \sum_{j=1}^{n} x_j^2 y_j \\ \sum_{j=1}^{n} x_j y_j \\ \sum_{j=1}^{n} y_j \end{bmatrix}$

$$\boxed{A^T A \, Z = A^T B} \longleftarrow \text{normal equations}$$

$$Z = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \longleftarrow \text{the vector of unknowns}$$

Example : Given the data

x	1	1.1	1.3	1.5	1.9	2.1
y	1.85	1.96	2.20	2.50	2.97	3.21

find the best fit by a quadratic polynomial $y = ax^2 + bx + c$.

Solution

Start with $\boxed{A^T A Z = A^T B}$ ✳

$Z = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ ← these are the numbers that we must find

? A , B. ⟿ available from the above table

$A = \begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \\ \vdots & \vdots & \vdots \\ x_6^2 & x_6 & 1 \end{bmatrix}$ ⟶ find A^T → find $A^T A$

$B = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_6 \end{bmatrix}$ ⟶ find $A^T B$

next : solve ✳

$$\begin{bmatrix} 42.8629 & 24.0230 & 14.17 \\ 24.0230 & 14.17 & 8.9 \\ 14.17 & 8.9 & 6 \end{bmatrix} \underbrace{}_{A^T A} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 38.4424 \\ 23 \\ 14.69 \end{bmatrix} \underbrace{}_{A^T B}$$

$\Rightarrow$ solve using Gaussian elimination .....