Approximation Theory. Discrete Least Squares Approximation.

Least Squares Error:

$$\boxed{E(a,b) = \sum_{i=1} \left[y_i - (ax_i + b) \right]^2} \qquad (1)$$

Necessary Condition to minimize (1).

$$\frac{\partial E}{\partial a}(a,b) = 0 \quad \text{and} \quad \frac{\partial E}{\partial b}(a,b) = 0.$$

$$\frac{\partial E}{\partial a} = 0 \Rightarrow \cancel{2} \sum_{i=1}^{m} \left(y_i - (ax_i + b) \right)(-x_i) = 0 \qquad (2)$$

$$\frac{\partial E}{\partial b} = 0 \Rightarrow \cancel{2} \sum_{i=1}^{m} \left(y_i - (ax_i + b) \right)(-1) = 0 \qquad (3)$$

From (2): $-\sum_{i=1}^{m} x_i y_i + a \sum_{i=1}^{m} x_i^2 + b \sum_{i=1}^{m} x_i = 0$

From (3): $-\sum_{i=1}^{m} y_i + a \sum_{i=1}^{m} x_i + \sum_{i=1}^{m} b = 0$

or equivalently,

$$\begin{cases} a \sum_{i=1}^{m} x_i^2 + b \sum_{i=1}^{m} x_i = \sum_{i=1}^{m} x_i y_i & (4) \\ a \sum_{i=1}^{m} x_i + b\, m = \sum_{i=1}^{m} y_i & (5) \end{cases}$$

Two linear equations – two unknowns "a" and "b".

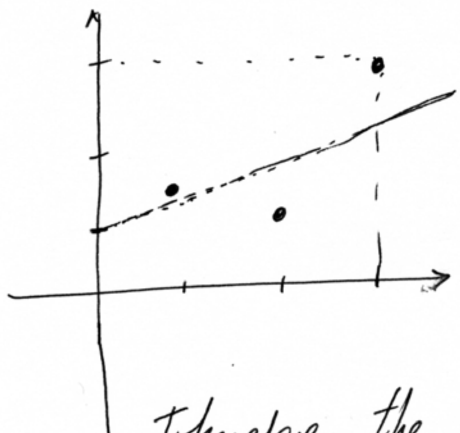
In matrix form:

$$\begin{bmatrix} \sum_{i=1}^{m} x_i^2 & \sum_{i=1}^{m} x_i \\ \sum_{i=1}^{m} x_i & m \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} x_i y_i \\ \sum_{i=1}^{m} y_i \end{bmatrix}$$

$\rightarrow$ These two equations are called "Normal Equations"

<u>Example:</u> Consider the three points

$$(x_1, y_1) = (1, 3/2), \quad (x_2, y_2) = (2, 1), \quad (x_3, y_3) = (3, 3)$$



	x_i	y_i	x_i^2	$x_i y_i$
	1	3/2	1	3/2
	2	1	4	2
	3	3	9	9
Σ	6	11/2	14	25/2

Therefore, the normal equations are

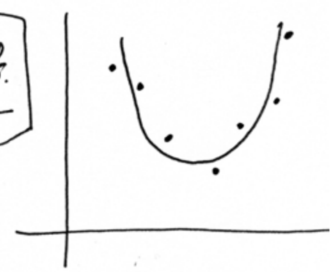
$$\begin{cases} 14a + 6b = \dfrac{25}{2} \\ 6a + 3b = \dfrac{11}{2} \end{cases}$$

Solns: $\boxed{a = 3/4}$ $\boxed{b = 1}$

So the least squares line approximating this data is given by

$$\boxed{y = \frac{3}{4}x + 1}$$

DATA POINTS DISTRIBUTION.



Look for general formulas in book.

Least squares Approx. with higher degree polynomials.

Parabola $\quad y = a_0 + a_1 x + a_2 x^2.$

$$E(a_0, a_1, a_2) = \sum_{i=1}^{m>3} \left[y_i - (a_0 + a_1 x + a_2 x^2) \right]^2$$

Necessary condition:

$$\frac{\partial E}{\partial a_i} = 0, \quad \text{for all } i = 0, 1, 2.$$

This means,

$$\frac{\partial E}{\partial a_0} = \cancel{2} \sum_{i=1}^{m} \left[y_i - (a_0 + a_1 x_i + a_2 x_i^2) \right](-1) = 0$$

$$\frac{\partial E}{\partial a_1} = \cancel{2} \sum_{i=1}^{m} \left[y_i - (a_0 + a_1 x_i + a_2 x_i^2) \right](-x_i) = 0$$

$$\frac{\partial E}{\partial a_2} = \cancel{2} \sum_{i=1}^{m} \left[y_i - (a_0 + a_1 x_i + a_2 x_i^2) \right](-x_i^2) = 0$$

Normal equations:

$$\begin{cases} a_0 m + a_1 \sum x_i + a_2 \sum x_i^2 = \sum y_i \\ a_0 \sum x_i + a_1 \sum x_i^2 + a_2 \sum x_i^3 = \sum x_i y_i \\ a_0 \sum x_i^2 + a_1 \sum x_i^3 + a_2 \sum x_i^4 = \sum x_i^2 y_i \end{cases}$$

In matrix form:

$$\begin{bmatrix} m & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \end{bmatrix}$$

In general for polynomials degree n.

$$P_n(x) = a_0 + a_1 x + \dots + a_n x^n$$

$$E(a_0, a_1, a_2, \dots, a_n) = \sum_{i=1}^{m} \left[y_i - P_n(x_i) \right]^2$$

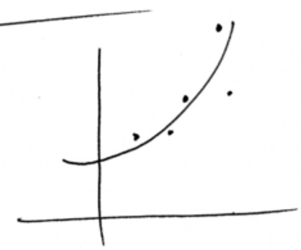
Normal equations are given in matrix form as

$$\overbrace{\begin{bmatrix} m & \sum x_i & \cdots & \sum x_i^n \\ \sum x_i & \sum x_i^2 & \cdots & \sum x_i^{n+1} \\ \vdots & & & \\ \sum x_i^n & \sum x_i^{n+1} & \cdots & \sum x_i^{2n} \end{bmatrix}}^{=A} \overbrace{\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}}^{=\vec{a}} = \overbrace{\begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \\ \vdots \\ \sum x_i^n y_i \end{bmatrix}}^{=\vec{b}}$$

<u>Theorem</u>: If $x_1, x_2, .., x_m$ are distinct with $m > n+1$ then the matrix A (above) is Symmetric and nonsingular. Therefore, there is a unique polynomial of degree "n" fitting the given data.

<u>Proof</u>.- Exercise # 14.

If data is exponentially related



use $\boxed{y = b e^{ax}}$

as a "fitting curve"

Least Squares error:

$$E(a,b) = \sum_{i=1}^{m} \left(y_i - b\, e^{a x_i} \right)^2$$

Normal equations:

$$\frac{\partial E}{\partial b} = 0 \Rightarrow 2 \sum_{i=1}^{m} \left(y_i - b\, e^{a x_i} \right) \left(-e^{a x_i} \right) = 0$$

$$\frac{\partial E}{\partial a} = 0 \Rightarrow 2 \sum_{i=1}^{m} \left(y_i - b\, e^{a x_i} \right) \left(-b x_i e^{a x_i} \right) = 0$$

Nonlinear system for "a" and "b".

$$\begin{cases} b\, e^{2 a x_i} - e^{a x_i} y_i = 0 \\ b^2 e^{2 a x_i} x_i - x_i e^{a x_i} y_i b = 0 \end{cases}$$

Hard to find Solutions!

Alternative: Consider logarithm of the "fitting curve"

$$y = b\, e^{a x} \Rightarrow \boxed{\ln y = \ln b + a x.}$$

then, normal equations are

$$\begin{cases} a\sum x_i^2 + \ln(b)\sum x_i = \sum x_i \ln y_i \\ a\sum x_i + \ln(b)\, m = \sum \ln y_i \end{cases}$$

Unknowns are "a" and "$\ln(b)$".

if $a = 2.5$ and $\ln(b) = 3. \Rightarrow b = e^3$

then $y = b e^{ax} = e^3 e^{2.5x} = e^{(2.5x+3)}$.