$$(x_1, y_1), \quad (x_2, y_2) \ldots, \quad (x_n, y_n)$$

$$\hat{y}(x) = a_1 + a_2 x \quad, \quad \hat{y}_i = \hat{y}(x_i)$$

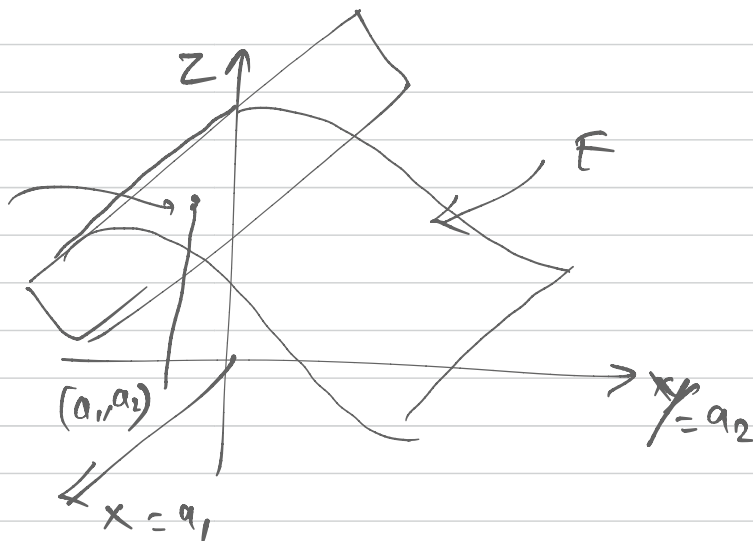$$(x_1, \hat{y}_1), \quad (x_2, \hat{y}_2) \ldots, \quad (x_n, \hat{y}_n)$$

$$E = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = E(a_1, a_2)$$

find a_1, a_2 s.t E is minimum

$$\frac{\partial E}{\partial a_1} = 0 \quad \Rightarrow \quad \boxed{a_1(n) + a_2\left(\sum_{i=1}^{n} x_i\right) = \left(\sum_{i=1}^{n} y_i\right)}$$

$$\frac{\partial E}{\partial a_2} = 0$$

$$\Rightarrow \boxed{a_1\left(\sum_{i=1}^{n} x_i\right) + a_2\left(\sum_{i=1}^{n} x_i^2\right) = \left(\sum_{i=1}^{n} x_i y_i\right)}$$

Z

E

(a_1, a_2)

$y = a_2$

$x = a_1$

$\Rightarrow Ja = b$,

$$a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad b = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}, \quad J = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix}$$

$$J = B^T B = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix}_{2 \times n} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}_{n \times 2}$$

$$= \begin{bmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{bmatrix}$$

$$\hat{y}(x) = a_1 + a_2 x + a_3 x^2$$

$$E = E(a_1, a_2, a_3) = \Sigma (y_i - \hat{y}_i)^2$$

$$\frac{\partial E}{\partial a_1} = 0 \quad \nearrow \quad a_1 (n) + a_2 (\Sigma x_i) + a_3 (\Sigma x_i^2) = \Sigma y_i$$

$$\frac{\partial E}{\partial a_2} = 0 \quad \nearrow \quad a_1 (\Sigma x_i) + a_2 (\Sigma x_i^2) + a_3 (\Sigma x_i^3) = \Sigma x_i y_i$$

$$\frac{\partial E}{\partial a_3} = 0 \quad \Rightarrow \quad a_1 (\Sigma x_i^2) + a_2 (\Sigma x_i^3) + a_3 (\Sigma x_i^4) = \Sigma x_i^2 y_i$$

$$J a = b \quad \nearrow \quad b = \begin{bmatrix} \Sigma y_i \\ \Sigma x_i y_i \\ \Sigma x_i^2 y_i \end{bmatrix} , \quad a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

$$J = \begin{bmatrix} n & \Sigma x_i & \Sigma x_i^2 \\ \Sigma x_i & \Sigma x_i^2 & \Sigma x_i^3 \\ \Sigma x_i^2 & \Sigma x_i^3 & \Sigma x_i^4 \end{bmatrix}$$

$$\hat{y}(x) = a_1 + a_2 x = [1 \quad x]_{1\times 2} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}_{2\times 1}$$

$$\hat{y}_1 = \hat{y}(x_1) = [1 \quad x_1] \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

$$\hat{y}_2 = \hat{y}(x_2) = [1 \quad x_2] \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

$$\vdots$$

$$\hat{y}_n = \hat{y}(x_n) = [1 \quad x_n] \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

$$\underline{\hat{y}} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix}_{n\times 1} = \underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}_{n\times 2}}_{B} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}_{2\times 1} \quad / \quad \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \underline{y}$$

$$a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

$$E = \sum_i (y_i - \hat{y}_i)^2$$

$$= (\underline{\hat{y}} - \underline{y})^T (\underline{\hat{y}} - \underline{y})$$

$$= (Ba - \underline{y})^T (Ba - \underline{y})$$

$$\underline{\hat{y}}^T \rightarrow 1\times n$$

$$\underline{y} \rightarrow n\times 1$$

$$\hat{y}(x) = a_1 + a_2 x = \begin{bmatrix} 1 & x \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \\ 1 & x_n \end{bmatrix}_{n \times 2} \qquad \text{n data point}$$

$$J = B^T B = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix}$$

$$b = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix} = B^T_{2 \times n} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n \times 1}$$

$$= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & & x_n \end{bmatrix}_{2 \times n} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$= \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

$$J a = b, \qquad J = B^T B, \qquad a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad b = B^T \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\hat{y} = a_1 + a_2 x + a_3 x^2 = \begin{bmatrix} 1 & x & x^2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} 1 & x & x^2 \end{bmatrix}}_{Z(x)}$$

$$B = \begin{bmatrix} - Z(x_1) - \\ - Z(x_2) - \\ \vdots \\ - Z(x_n) - \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & & \\ 1 & x_n & x_n^2 \end{bmatrix}$$

$$J a = b, \qquad J = B^T B$$

$$b = B^T \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

General linear regression least square method

$$\hat{y}(x) = \underset{1 \times m}{Z(x)} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}_{m \times 1} \Longrightarrow$$

$$(x_i, y_i) \quad i = 1, 2, \ldots, n$$

$$J a = b$$

$$J = B^T B$$

$$b = B^T \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$Z(x) = \begin{bmatrix} z_1(x), & z_2(x), & \cdots & , z_m(x) \end{bmatrix}_{1 \times m}$$

$\downarrow$

basis for a curve

$$B = \begin{bmatrix} - Z(x_1) - \\ - Z(x_2) - \\ \vdots \\ - Z(x_n) - \end{bmatrix}$$

(i) straight line $m = 2$

$$Z(x) = [1, x]$$

$$Z_1(x) = 1$$

$$Z_2(x) = x$$
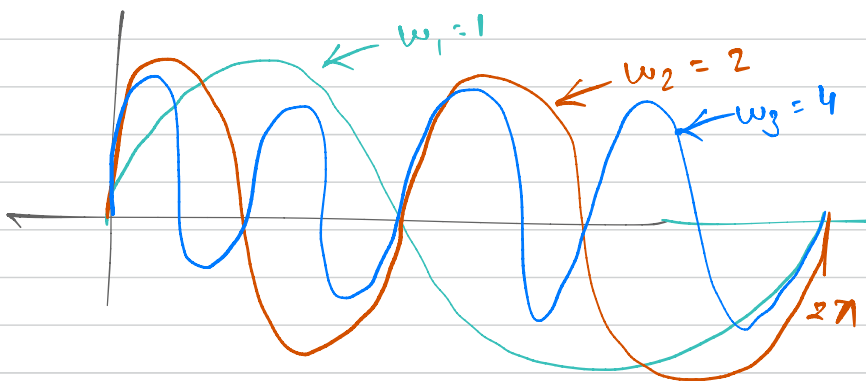
(ii) quadratic function $m = 3$

$$Z(x) = [1, x, x^2]$$

(iii) $(m-1)^{th}$ order polynomial

$$Z(x) = [1, x, x^2, \ldots, x^{m-1}]_{1 \times m}$$

(iv) Fourier basis
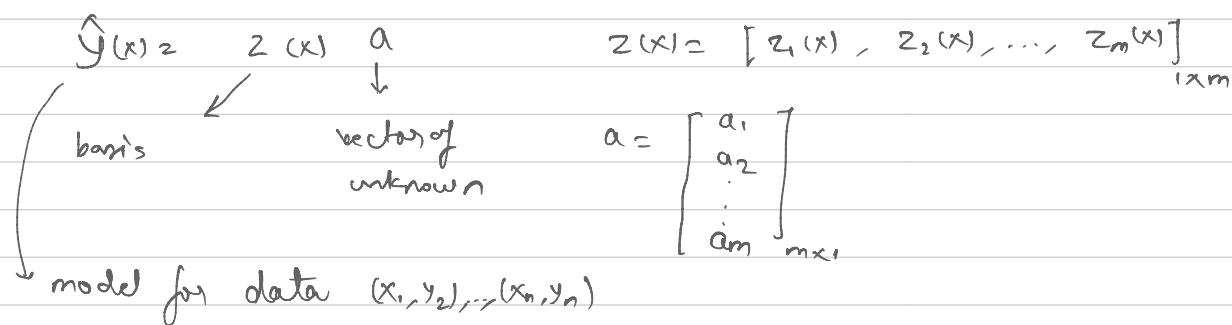
$$Z(x) = [1, \sin(\omega_1 x), \sin(\omega_2 x), \ldots, \sin(\omega_{m-1} x)]_{1 \times m}$$

$\leftarrow \omega_1 = 1$

$\leftarrow \omega_2 = 2$

$\leftarrow \omega_3 = 4$

2π

$\omega_2 = 2$

$\omega_3 = 4$

$\hat{y}(x) = z(x)\, a$

$z(x) = [z_1(x),\ z_2(x),\ \dots,\ z_m(x)]_{1 \times m}$

basis $\qquad$ vector of unknown

$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}_{m \times 1}$

model for data $(x_1, y_2), \dots, (x_n, y_n)$

let

$$\underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n \times 1}, \qquad \hat{\underline{y}} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix}_{n \times 1} = \begin{bmatrix} z_1(x_1) & z_2(x_1) & \cdots & z_m(x_1) \\ z_1(x_2) & z_2(x_2) & \cdots & z_m(x_2) \\ \vdots & & & \\ z_1(x_n) & z_2(x_n) & \cdots & z_m(x_n) \end{bmatrix}_{n \times m} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}_{m \times 1}$$

$$\Rightarrow \hat{\underline{y}} = B\, a \qquad\qquad B = \begin{bmatrix} - & z(x_1) & - \\ - & z(x_2) & - \\ & \vdots & \\ - & z(x_n) & - \end{bmatrix}_{n \times m}$$

Squared error $\qquad E = \displaystyle\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

$$= (\underline{y} - \hat{\underline{y}})^T (\underline{y} - \hat{\underline{y}})$$

for $j = 1, 2, \dots, m$

$$\frac{\partial E}{\partial a_j} = \sum_{i=1}^{n} \frac{\partial}{\partial a_j} (y_i - \hat{y}_i)^2$$

$$= \sum_{i=1}^{n} 2(y_i - \hat{y}_i)\left(-\frac{\partial \hat{y}_i}{\partial a_j}\right)$$

where $\dfrac{\partial y_i}{\partial a_j} = 0$, & $\dfrac{\partial \hat{y}_i}{\partial a_j} = \dfrac{\partial}{\partial a_j}\Big[z(x_i)\, a\Big]$

$$= \frac{\partial}{\partial a_j}\left\{ [z_1(x_i),\ z_2(x_i),\ \dots,\ z_m(x_i)] \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \right\}$$

$$= \frac{\partial}{\partial a_j} \left\{ z_1(x_i) a_1 + z_2(x_i) a_2 + z_3(x_i) a_3 + \cdots + z_m(x_i) a_m \right\}$$

$$= z_j(x_i)$$

for $j = 1, 2, \ldots, m$

$$\frac{\partial E}{\partial a_j} = \sum_{i=1}^{n} 2 (y_i - \hat{y}_i) \, z_j(x_i)$$

$$= 2 \left[\sum_{i=1}^{n} y_i \, z_j(x_i) - \sum_{i=1}^{n} \hat{y}_i \, z_j(x_i) \right]$$

$$\rightarrow \quad \frac{\partial E}{\partial a_j} = 0 \quad \rightarrow \quad \sum_{i=1}^{n} \hat{y}_i \, z_j(x_i) = \sum_{i=1}^{n} y_i \, z_j(x_i)$$

$$\because \quad \hat{y}_i = z(x_i) \, a$$

$$\therefore \quad \hat{y}_i \, z_j(x_i) = z_j(x_i) \, z(x_i) \, a$$

$$\rightarrow \quad \sum_{i=1}^{n} \hat{y}_i \, z_j(x_i) = \sum_{i=1}^{n} z_j(x_i) \, z(x_i) \, a$$

$$= z_j(x_1) \, z(x_1) \, a + z_j(x_2) \, z(x_2) \, a + \cdots + z_j(x_n) \, z(x_n) a$$

$$\because \quad z(x_i) = \left[z_1(x_i), z_2(x_i), \ldots, z_m(x_i) \right]$$

$$= \left(z_j(x_1) \, z(x_1) \right) a + \left(z_j(x_2) \, z(x_2) \right) a + \cdots$$

$$+ \left(z_j(x_n) \, z(x_n) \right) a$$

$$= \left(\left[z_j(x_1), z_j(x_2), \ldots, z_j(x_n) \right] \underset{1 \times n}{} \begin{bmatrix} z(x_1) \\ z(x_2) \\ \vdots \\ z(x_n) \end{bmatrix}_{n \times m} \right) a_{m \times 1}$$

$$\Downarrow$$

$$B_{n \times m}$$

$$= \left[z_j(x_1), \ldots, z_j(x_n) \right] B \, a$$

$$\stackrel{!}{=} \quad [z_j(x_1), \ldots, z_j(x_n)] \; \beta \, a \; = \; \sum_{i=1}^{n} z_j(x_i) y_i \quad \Rightarrow$$

$$\underbrace{\phantom{\sum_{i=1}^{n} z_j(x_i) y_i}}_{b_j}, \quad j = 1, 2, \ldots, m$$

Stacking above gives , $j = 1, 2, \ldots, m$,

$$\begin{bmatrix} z_1(x_1) & z_1(x_2) & \cdots & z_1(x_n) \\ z_2(x_1) & z_2(x_1) & \cdots & z_2(x_n) \\ \vdots & & & \\ z_m(x_1) & z_m(x_1) & \cdots & z_m(x_n) \end{bmatrix} \beta \, a = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

$$\underbrace{\phantom{\begin{bmatrix} z_1(x_1) & z_1(x_2) & \cdots & z_1(x_n) \end{bmatrix}}}_{\text{it is } B^T \; (\text{check!})}$$

$$\Rightarrow \qquad B^T B \, a = b$$

Now $\quad b_j = \sum_{i=1}^{n} z_j(x_i) y_i$

$$= \quad [z_j(x_1), z_j(x_2) \ldots, z_j(x_n)] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \longrightarrow \underline{y}$$

$$\therefore \quad \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} z_1(x_1) & z_1(x_2) & \cdots & z_1(x_n) \\ \vdots & & & \\ z_m(x_1) & z_m(x_2) & \cdots & z_m(x_n) \end{bmatrix} \underline{y}$$

$$\underbrace{\phantom{\begin{bmatrix} z_1(x_1) & z_1(x_2) & \cdots & z_1(x_n) \end{bmatrix}}}_{\text{it is } B^T}$$

$$\Rightarrow \qquad b = B^T \underline{y}$$

__Then__ solve for $a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}$ using $\qquad \underbrace{B^T B}_{J \text{ matrix}} a = \underbrace{B^T \underline{y}}_{b \text{ vector}}$

- **Linear regression with multiple variables**

So far we considered data of type $(x_1, y_1), \ldots, (x_n, y_n)$

where x_i, y_i are scalars.

we could have data of the form

$$(x_1, y_1), \quad (x_2, y_2), \quad \ldots \quad, \quad (x_n, y_n)$$

where y_i scalar

$$x_i = \begin{bmatrix} x_i^1 \\ x_i^2 \\ \vdots \\ x_i^m \end{bmatrix} \quad \text{vector of } m \text{ components.}$$
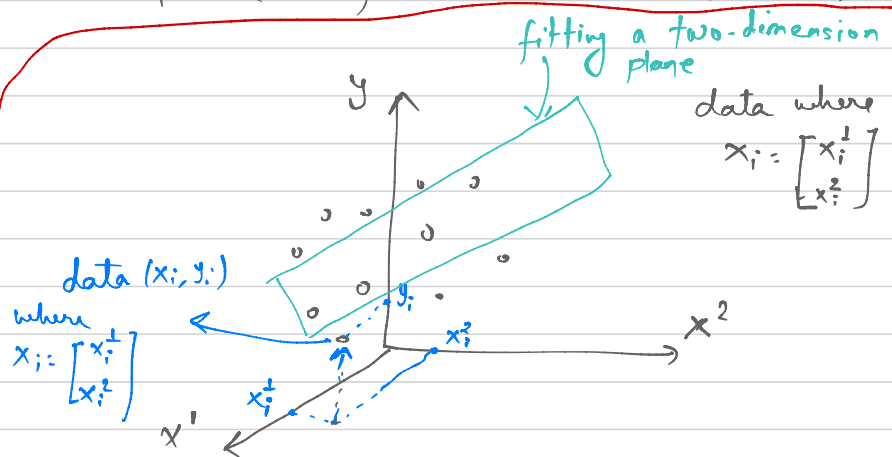
To fit this "m-dimensional" data using an m-dimensional plane in $m+1$-dimensional space (m from x_i and 1 from y_i)

Take the model (m-dimensional plane, i.e., linear function of m arguments)

$$\hat{y} = \hat{y}(x) = a_1 + a_2 x^1 + a_3 x^2 + \ldots + a_{m+1} x^m$$

where

$$X = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^m \end{bmatrix}$$

fitting a two-dimension plane

data where $x_i = \begin{bmatrix} x_i^1 \\ x_i^2 \end{bmatrix}$

data (x_i, y_i)
where $x_i = \begin{bmatrix} x_i^1 \\ x_i^2 \end{bmatrix}$

in this case, $m = 2$ and plot of data is in $m+1 = 3$ dimension.

$\downarrow$

fit 2-dimension plane (linear in x^1 and x^2)

Defining basis vector $Z = Z^{(x)}$ as

$$Z(x) = \begin{bmatrix} 1, & x^1, & x^2, & \ldots & , x^m \end{bmatrix}_{1 \times (m+1)} \quad \longrightarrow \quad \text{linear in } m \text{ arguments}$$

$$\downarrow$$

where $X = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^m \end{bmatrix}$
$\qquad\qquad$ extension of
$\qquad\qquad$ linear curve

Our general linear regression formula still holds

find $\quad a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{m+1} \end{bmatrix}_{(m+1) \times 1}$ $\quad$ s.t.

$$J a = b$$

where

$$J_{(m+1) \times (m+1)} = B^T B, \quad b_{(m+1) \times 1} = B^T y, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n \times 1}$$

$\underline{\underline{\ell}}$

$$B_{n \times (m+1)} = \begin{bmatrix} \underline{\qquad} Z(X_1) \underline{\qquad} \\ \underline{\qquad} Z(X_2) \underline{\qquad} \\ \cdot \\ \cdot \\ \underline{\qquad} Z(X_n) \underline{\qquad} \end{bmatrix}$$
$\quad$ where $X_i = \begin{bmatrix} x_i^1 \\ x_i^2 \\ \vdots \\ x_i^m \end{bmatrix}_{m \times 1}$

$$= \begin{bmatrix} 1 & x_1^1 & x_1^2 & \ldots & x_1^m \\ 1 & x_2^1 & x_2^2 & \ldots & x_2^m \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n^1 & x_n^2 & \ldots & x_n^m \end{bmatrix}$$

We could consider also polynomial curve in m-arguments

E.g. take $m = 2$ and consider 2nd order polynomial of two variables

$$\hat{y}(x) = \underbrace{\begin{bmatrix} 1 & x^1 & x^2 & (x^1)^2 & (x^2)^2 & x^1 x^2 \end{bmatrix}}_{Z(x)} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix}$$

Then $Ja = b$, where $a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_6 \end{bmatrix}$, $b = B^T \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, $J = B^T B$

and

$$B = \begin{bmatrix} \text{------} & Z(x_1) & \text{------} \\ \text{------} & Z(x_2) & \text{------} \\ & \cdot & \\ & \cdot & \\ \text{------} & Z(x_n) & \text{------} \end{bmatrix}$$

- ## Nonlinear regression

Consider following model for n data $(x_1, y_1), \ldots, (x_n, y_n)$

(assume x_i, y_i are scalars)

$$\hat{y}(x) = a_1 + a_2 \sin(a_3 x) \qquad \text{where } a_1, a_2, a_3 \text{ need to be } \underline{\text{computed}}.$$

This is an example of $\underline{\text{nonlinear regression}}$

This leads to minimization of nonlinear function

Total error $\quad E = E(a_1, a_2, a_3)$

$$= \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \qquad , \quad \hat{y}_i = \hat{y}(x_i)$$

find a_1, a_2, a_3 s.t $\quad E(a_1, a_2, a_3)$ is minimum

$$\Downarrow$$

Solve nonlinear equations

$$\frac{\partial E}{\partial a_1} = 0 \quad \Rightarrow \quad \sum 2(y_i - \hat{y}_i)(-1) = 0 \Rightarrow \sum \left[y_i - (a_1 + a_2 \sin(a_3 x_i)) \right] = 0$$

$$\frac{\partial E}{\partial a_2} = 0 \quad \Rightarrow \quad \sum 2(y_i - \hat{y}_i)(-\sin(a_3 x_i)) = 0 \Rightarrow \sum \left[y_i - (a_1 + a_2 \sin(a_3 x_i)) \right] \sin(a_3 x_i) = 0$$

$$\frac{\partial E}{\partial a_3} = 0 \quad \Rightarrow \quad \sum 2(y_i - \hat{y}_i)\underbrace{\left(-a_1 \cos(a_1 x_i) x_i \right)}_{\text{check !!}} = 0$$

$$\Rightarrow \quad \sum \left[y_i - (a_1 + a_2 \sin(a_3 x_i)) \right] a_1 \cos(a_3 x_i) x_i = 0$$

Read Section 15.5 and example 15.5

"fminsearch" $\Downarrow$ MATLAB method to find minima of nonlinear function.