Lecture 40

- Linear regression in one-dimensional setting

$\Rightarrow$ (x^1, y^1), (x^2, y^2), . . . , (x^n, y^n) n-pairs of data

x^i = independent variable (parameter) associated to data y^i

$\Rightarrow$ find a model $\hat{y} = \hat{y}(x)$ for the data

is scalar hence $\hat{y}$ is one-dimensional function

$\Rightarrow$ linear regression

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

where

$$z(x) = [z_1(x), z_2(x), \ldots, z_m(x)] \leftarrow \text{"known"}$$

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \leftarrow \text{"unknown"}$$

Examples: (i) $m = 2$ and $\hat{y}$ to be line

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

(ii) $m = 2$ and $\hat{y}$ to be quadratic function

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

(iii) $m = 3$ and $\hat{y}$ to be cubic function

$$z(x) = [1, x, x^2, x^3]$$

(iv) $m = 3$ and $\hat{y}$ to be exponential - like function

$z(x) = [\exp[\lambda_1 x], \exp[\lambda_2 x], \exp[\lambda_3 x]]$

and $\lambda_1, \lambda_2, \lambda_3$ are "known" and "fixed"

$\lambda_1 \neq \lambda_2 \neq \lambda_3$

(v) $m = 4$ and $\hat{y}$ to be sinusoidal - like function

$z(x) = [\sin(\omega_1 x), \sin(\omega_2 x), \sin(\omega_3 x), \sin(\omega_4 x)]$

$\omega_1, \omega_2, \omega_3, \omega_4$ are "known" and "fixed"

$\omega_1 \neq \omega_2 \neq \omega_3 \neq \omega_4$

Solving linear regression

- Error = squared error

$$E = E(a) = \frac{1}{n} \sum_{i=1}^{n} (y^i - \hat{y}(x^i))^2$$

- Problem:

find "a" such that $E(a)$ is minimum

$$\Updownarrow$$

this is shown in Lecture 26

$$Ja = b$$

where

$$J = B^T B \longrightarrow m \times m \text{ matrix}$$

$$b = B^T \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^n \end{bmatrix} = B^T y \longrightarrow m \times 1 \text{ vector}$$

$$y = \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^n \end{bmatrix}$$

$$B = \begin{bmatrix} \text{---} & z(x^1) & \text{---} \\ \text{---} & z(x^2) & \text{---} \\ & \vdots & \\ \text{---} & z(x^n) & \text{---} \end{bmatrix} \longrightarrow n \times m \text{ matrix}$$

$$= \begin{bmatrix} z_1(x^1) & z_2(x^1) & \cdots & z_m(x^1) \\ z_1(x^2) & z_2(x^2) & \cdots & z_m(x^2) \\ \vdots & & & \\ z_1(x^n) & z_2(x^n) & \cdots & z_m(x^n) \end{bmatrix}$$

- Linear regression in multiple dimensional setting

Data: temperature at different points on plate

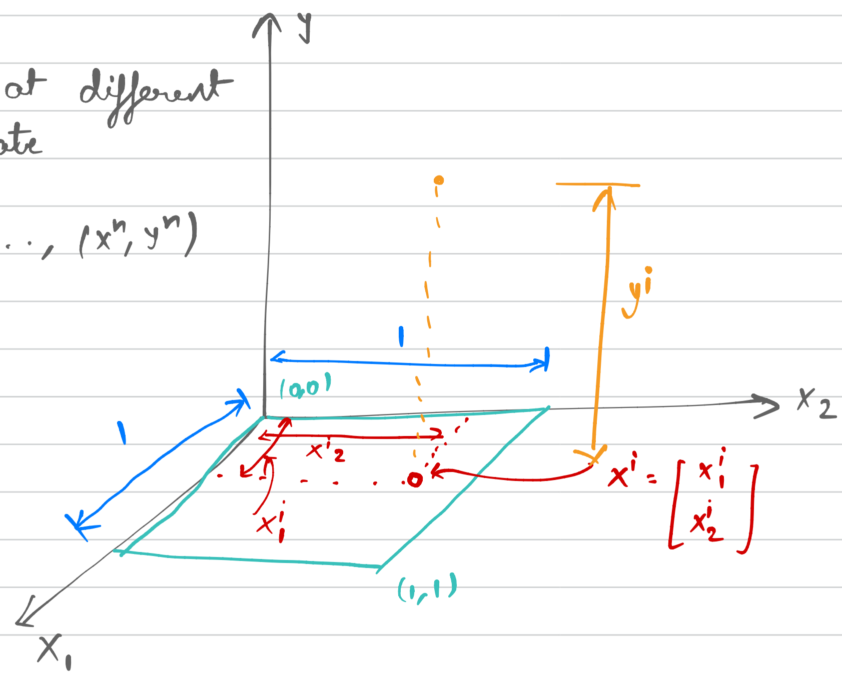
$(x^1, y^1), (x^2, y^2), \ldots, (x^n, y^n)$

for i^{th} data

y^i = temperature at point x^i

x^i = coordinate of ith point

$$= \begin{bmatrix} x_1^i \\ x_2^i \end{bmatrix}$$



Goal: find a model $\hat{y} = \hat{y}(x)$ for temperature data

x is a vector of two elements & hence $\hat{y}$ is a multi-dimensional function

Linear regression and using linear function, model $\hat{y}$ can be written as

$$\hat{y}(x) = a_1 + a_2 x_1 + a_3 x_2$$

$$= z(x)\, a$$

$$z(x) = [1, x_1, x_2], \quad a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

using least square method, we have

"Problem" find a s.t. $E(a) = \frac{1}{n} \sum_{i=1}^{n} (y^i - \hat{y}(x^i))^2$ is minimum

$\Updownarrow$

find a s.t. $Ja = b$

where $J = B^T B$, $b = B^T y$, $y = \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^n \end{bmatrix}$

$$B = \begin{bmatrix} \text{---} & Z(x^1) & \text{---} \\ \text{---} & Z(x^2) & \text{---} \\ & \vdots & \\ \text{---} & Z(x^n) & \text{---} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & x_1^1 & x_2^1 \\ 1 & x_1^2 & x_2^2 \\ 1 & x_1^3 & x_2^3 \\ \vdots & & \\ 1 & x_1^n & x_2^n \end{bmatrix}$$

- Linear regression using linear function in K-dimension

$(x^1, y^1), (x^2, y^2), \ldots, (x^n, y^n)$ n-data set

for each i $\quad x^i = \begin{bmatrix} x_1^i \\ x_2^i \\ \vdots \\ x_K^i \end{bmatrix}$

model

$$\hat{y}(x) = a_1 + a_2 x_1 + a_3 x_2 + \ldots + a_{K+1} x_K$$

$$= Z(x)\, a$$

$$Z(x) = [1, x_1, x_2, \ldots, x_K]$$

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{K+1} \end{bmatrix}$$

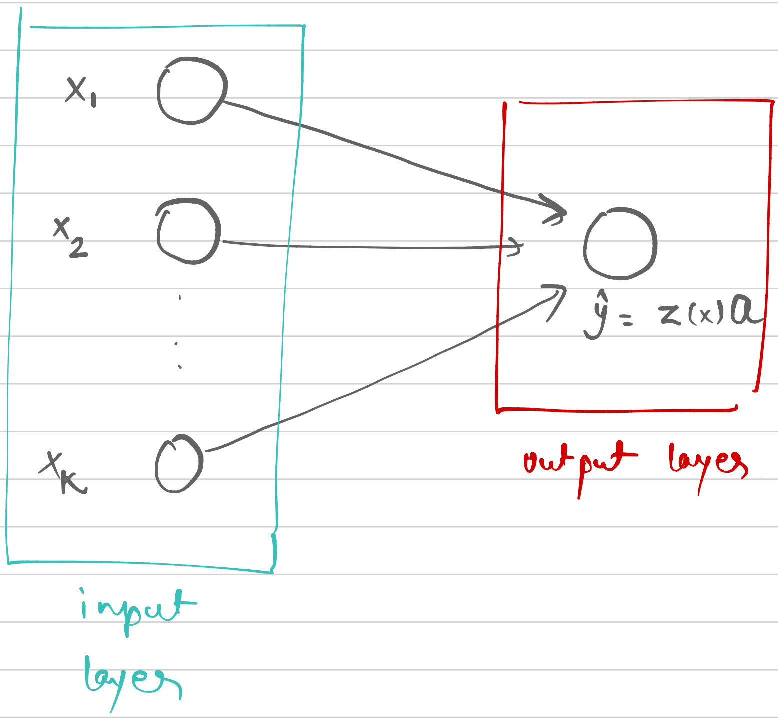
→ Solve $Ja = b$ (equivalent to minimizing squared error)

$$J = B^T B, \quad b = B^T y, \quad y = \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^n \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & x_1^1 & x_2^1 & \cdots & x_K^1 \\ 1 & x_1^2 & x_2^2 & \cdots & x_K^2 \\ \vdots & & & & \\ 1 & x_1^n & x_2^n & \cdots & x_K^n \end{bmatrix}$$

Graphical representation of linear regression in k-dimension

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \end{bmatrix}$$



- Nonlinear relation between input x and output y

x,

$$\hat{y} = f(x; \theta)$$

input

are model parameters

$$= f^{(3)}\left(f^{(2)}\left(f^{(1)}(x; \theta^{(1)}); \theta^{(2)}\right); \theta^{(3)}\right)$$

$$\boxed{f^{(3)} \circ f^{(2)} \circ f^{(1)} \circ id(x) = \hat{y}}$$