SC201 Lecture 2

Define Model (h(θ) = θx) h(θ) = _____

↓

Define Loss Function (L2 Loss) loss = _____

↓

Define Cost Function J J = _____

↓

Find the best parameter (θ) for i in range(iteration):

$$\theta = \theta - \boxed{\alpha \frac{dJ}{d\theta}}$$

$$\frac{dJ}{d\theta} = \frac{d}{d\theta}\left(\frac{\sum_{i=1}^{m} (\theta xi - yi)^2}{2m} \right)$$

$$\frac{dJ}{d\theta} = \frac{d}{d\theta}\left(\frac{\sum_{i=1}^{m} (yi - \theta xi)^2}{2m} \right)$$

Fine the best parameter (θ) → Training

① Initialization

② if $\frac{dJ(\theta)}{d\theta} > 0$, θ ↓

elif $\frac{dJ(\theta)}{d\theta} < 0$, θ ↑

elif $\frac{dJ(\theta)}{d\theta} == 0$, θ best

③ Gradient Descent

$\theta = \theta - \alpha\frac{dJ}{d\theta}$

④ Iteration

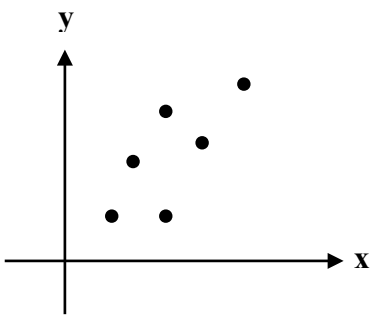
① <Step size (Learning Rate) α >

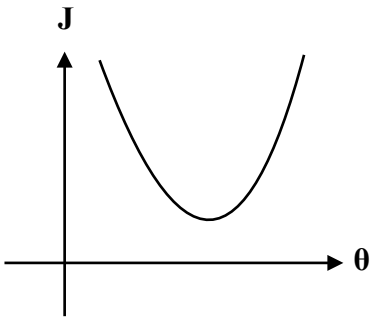
usually _____

② <Iteration >

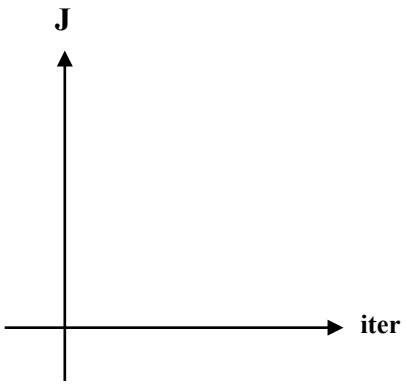
for i in range(iteration):

$\theta = \theta - \alpha\frac{dJ}{d\theta}$



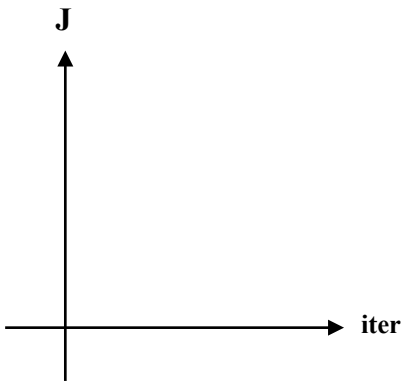


③ <Cost v.s. Iteration Plateau >



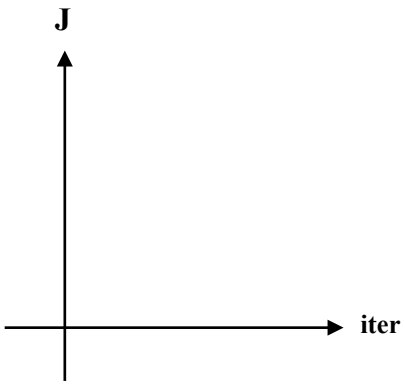
iter _____

α _____



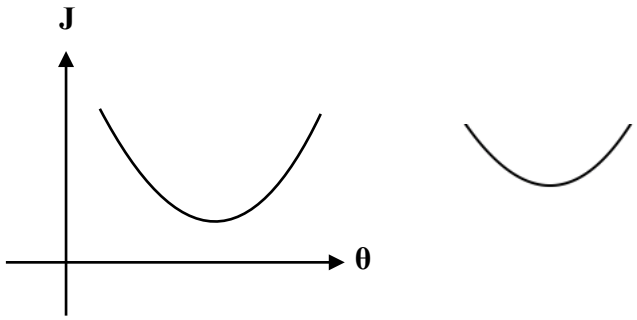
iter _____

α _____



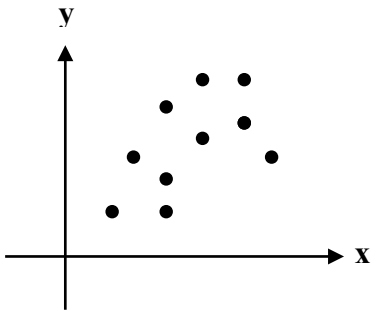
α _____

如果for i in range (100) 會超過最小值嗎？



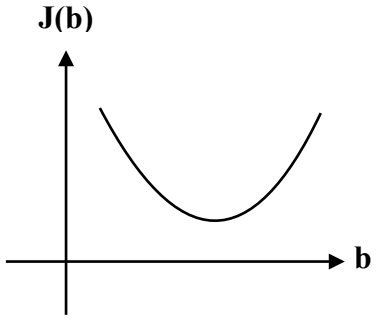
如果Model不一定過(0,0)?

new_model ⇛ _____



_____也是___的函數！

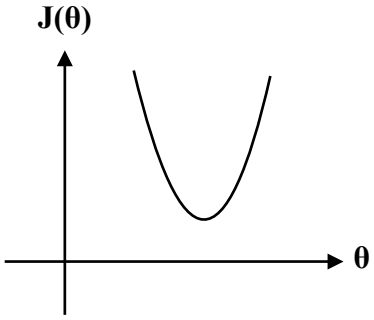
< When θ is fixed >



b = _____

< dJ_db >

< When b is fixed >



θ = _____

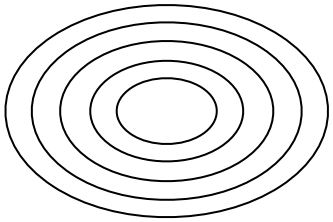
< dJ_dθ >

Non-Steepest Descent

for i in range(iter):

θ = _____

b = _____



- _____ update
- _____ to converge

Steepest Descent

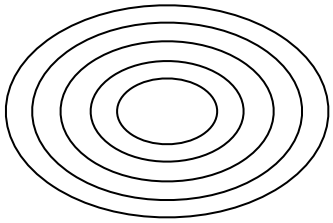
for i in range(iter):

θ_temp = _____

b_temp = _____

θ = _____

b = _____



- _____ update
- _____ to converge

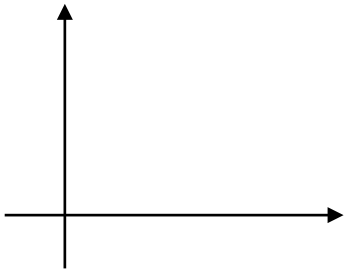
Polyminal Features

- raising existing features to an exponent

<degree1>
linear function

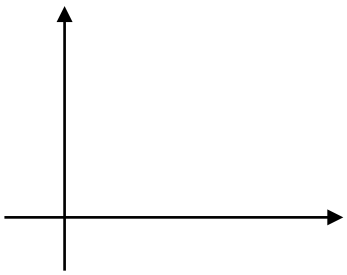
$h(\theta,b) = \theta x + b$

_____
(Model we chose)



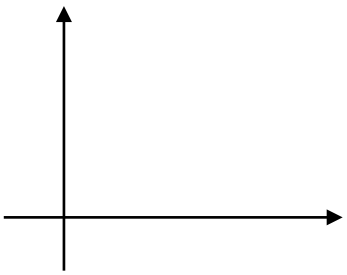
<degree2>
quadratic function

$h(\theta', \theta, b) = \theta' x^2 + \theta x + b$



<degree3>
cubic function

$h(\theta'', \theta', \theta, b) = \theta'' x^3 + \theta' x^2 + \theta x + b$



$$J = \frac{1}{2m} \sum_{i=1}^{m} (\theta' xi^2 + \theta xi + b - yi)^2$$

<dJ_dθ'>	<dJ_dθ>	<dJ_db>

Normalizationn

$$X' = \frac{\quad}{\quad}$$