SC201 Lecture 1

Machine Learning

data(i)	area(x)	\$(y)
data(1)	1	1
data(2)	2	2
data(3)	3	4
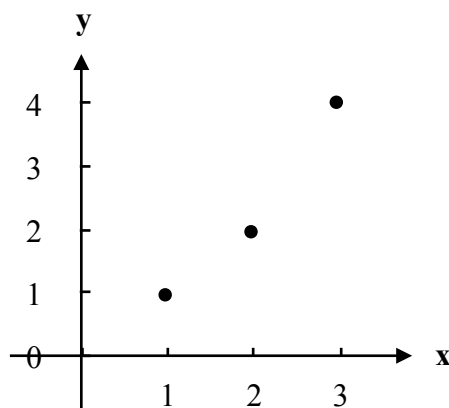
<Model 1>   y = x
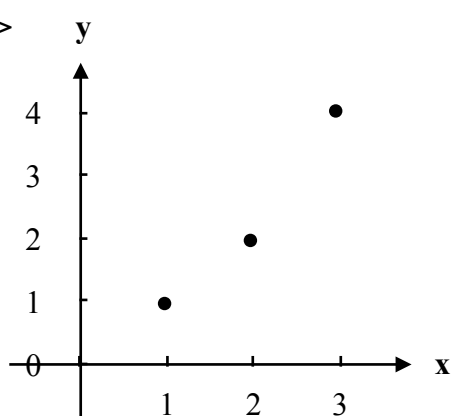
<Model 2>   y = 2
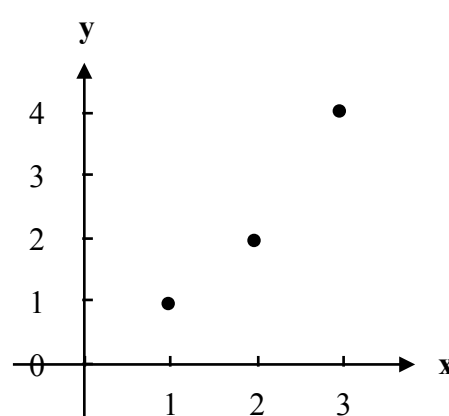

How do we know <model 1> y=x is better?

Loss Function

- How does your model perform on _____

1. 假設 loss(data(i)) = $y_i - y_i'$

    <Model 1>   y = x   <Model 2>   y = 2

2. **L1 loss function** $|y_i - y_i'|$

    <Model 1>   y = x   <Model 2>   y = 2

3. **L2 loss function** $(y_i - y_i')^2$

    <Model 1>   y = x   <Model 2>   y = 2

- To evaluate a model, we cannot do it on 1 data point (1 loss)!
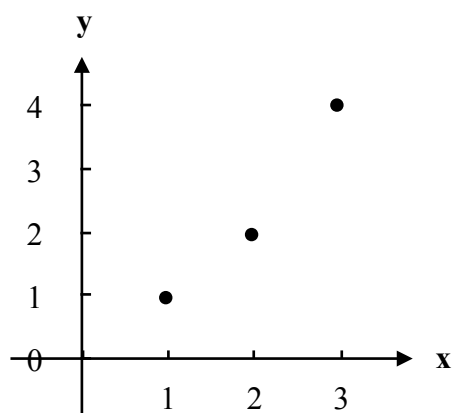
  Instead, we need _____

Cost Function

$$J = \underline{\hspace{6cm}}$$

(Note: where m is number of data)

cost-function.ipynb → **Jupyter notebook**

1. Web-based IDE
2. _____(not def main())
3. _____
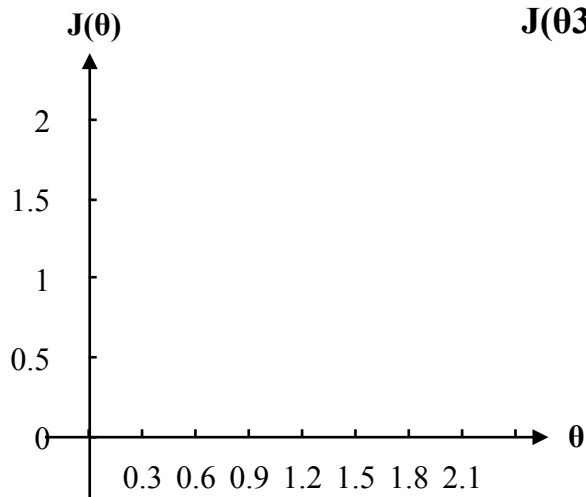4. 執行cell: _____ + _____

How do we find the <u>best line</u> by learning?

假設最好的model為 **h(θ) = θx**
→ y = mx 必過原點

- 哪個model最好?

  J(θ1) = _____

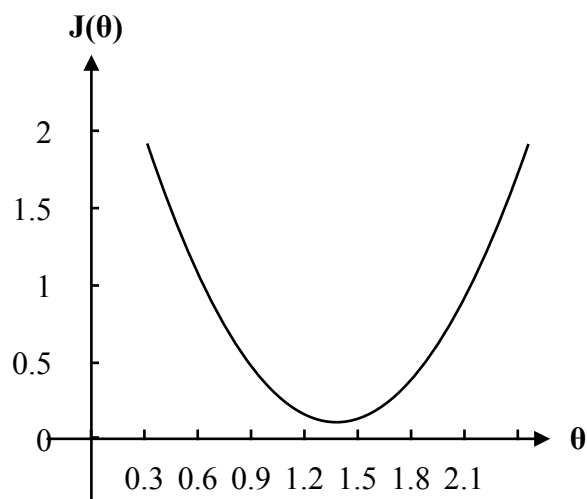 **J(θ2) =** _____

  J(θ3) = _____

What is this? →
_____

- 每個斜率都有一個cost

  _____ 是 **θ** 的函數

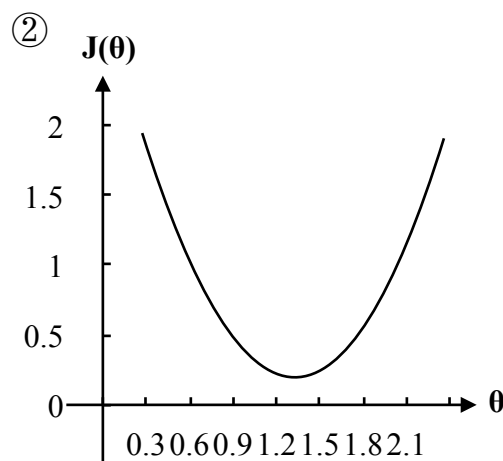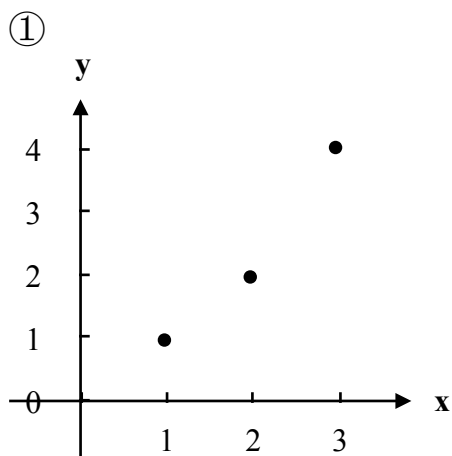- 當 θ 從 2 → 1/2 : Cost 從 ___ → ___ → ___
  (θ1)　　(θ3)

Linear Regression

- A way of finding the _____ of a function.

$$\begin{cases} \bullet \quad \text{If } \frac{dJ(\theta i)}{d\theta} > 0 \;\rightarrow\; \text{the min is at the } \underline{\qquad\qquad} \\[2em] \bullet \quad \text{If } \frac{dJ(\theta i)}{d\theta} < 0 \;\rightarrow\; \text{the min is at the } \underline{\qquad\qquad} \end{cases}$$

Training

① _____ (隨便選θi)

② if $\frac{dJ(\theta i)}{d\theta} < 0$, _____ θi ;

elif $\frac{dJ(\theta i)}{d\theta} > 0$, _____ θi ;

elif $\frac{dJ(\theta i)}{d\theta} == 0$, Congrats! You've found the min!

④ _____

③ θi = θi - ___ $\frac{dJ(\theta i)}{d\theta}$ (_____)

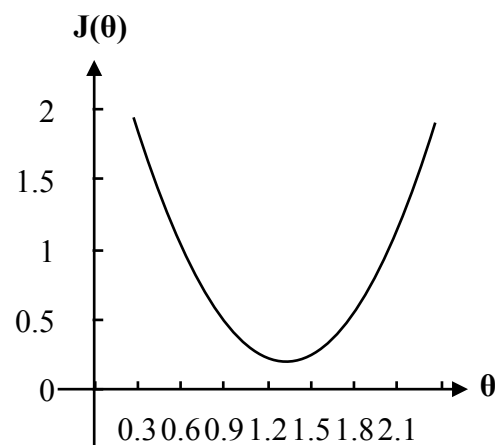①

② $J(\theta)$

③ θi = θi - α$\frac{dJ}{d\theta}$(θi)

Gradiant Descent
(絕對要做不只一次)

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

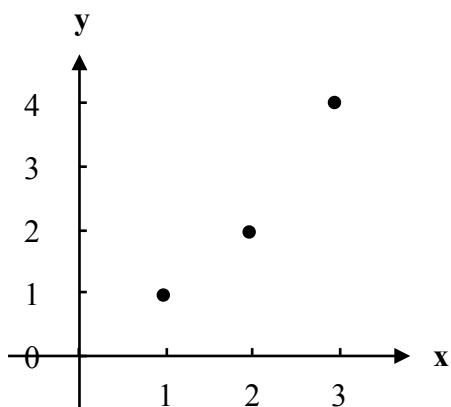- 當 θi 離 min 愈 ____ , $\left|\frac{dJ(\theta i)}{d\theta}\right|$ 愈 ____ , $\alpha\left|\frac{dJ}{d\theta}(\theta i)\right|$ 愈 ____ ➡ 更新愈 ____

- 當 θi 離 min 愈 ____ , $\left|\frac{dJ(\theta i)}{d\theta}\right|$ 愈 ____ , $\alpha\left|\frac{dJ}{d\theta}(\theta i)\right|$ 愈 ____ ➡ 更新愈 ____

Learninng Rate α

< α 太小 >

< α 太大 >

J上升 (只有α 太 ____ 的可能)

- 適中的α電腦無法跟你說，α 稱為 _____

Derivatives

f(x) = x²	f(x) = 3x²	g(x) = 2x	k(x) = 2
$\dfrac{df(x)}{dx} = f'(x) =$	$\dfrac{df(x)}{dx} = f'(x) =$	$\dfrac{dg(x)}{dx} = g'(x) =$	$\dfrac{dg(x)}{dx} = k'(x) =$

- $\dfrac{d\log x}{dx} =$

- $\dfrac{de^x}{dx} =$

Chain Rule

f(x) = (1-x)²

$\dfrac{df(x)}{dx} = ?$

$\boxed{\dfrac{dJ(\theta)}{d\theta}}$