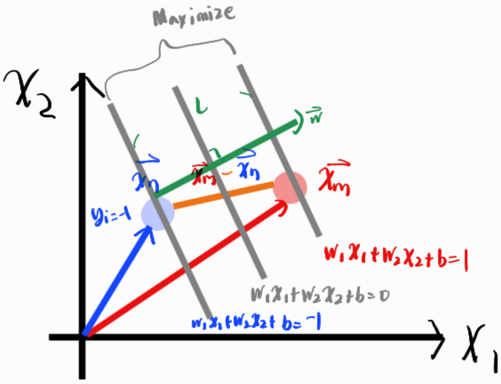
HW 1

1. SVM

$\vec{w} = (w_1, w_2)$ $\quad \vec{x_m} = (x_{1m}, x_{2m})$ $\quad \vec{x_n} = (x_{1n}, x_{2n})$

$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2$



① $w_1 x_{1m} + w_2 x_{2m} + b = 1$

② $w_1 x_{1n} + w_2 x_{2n} + b = -1$

③ ⇒ ① − ② ⇒ $w_1 (\overset{\Delta x_1}{x_{1m} - x_{1n}}) + w_2 (\overset{\Delta x_2}{x_{2m} - x_{2n}}) = 2$

$\vec{w} \cdot (\Delta x) = 2$

$\Rightarrow \vec{w} \cdot (\vec{x_m} - \vec{x_n}) = 2$

④ $\vec{w} \cdot (\vec{x_m} - \vec{x_n}) = 2 \Rightarrow |\vec{w}|\,|\vec{x_m} - \vec{x_n}| \cos\theta = 2$

$\Rightarrow$ Maximize $= L = \frac{2}{|\vec{w}|} = |\vec{x_m} - \vec{x_n}| \times \cos\theta$

正射影 $(\vec{x_m} - \vec{x_n})$ 在 $\vec{w}$ 的投影 L

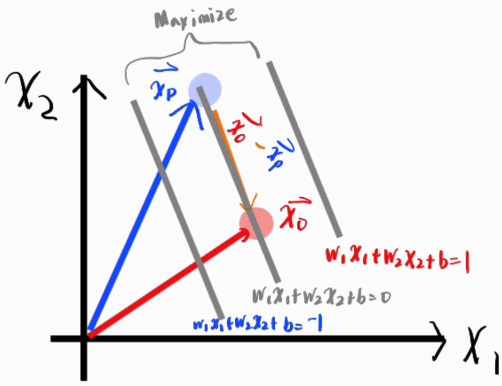
$L = (\overset{\Delta x}{\vec{x_m} - \vec{x_n}}) \cos\theta = \frac{|\Delta x \cdot \vec{w}|}{|w|} =$

$\Delta w \cdot \vec{w} = |\Delta w|\,|\vec{w}| \cos\theta$

$y_i = 1 \begin{cases} y_i * (\vec{w} \cdot \vec{x_i} + b) \geq 1 \\ y_i * (\vec{w} \cdot \vec{x_i} + b) \leq -1 \end{cases}$ $y_i = -1$

$\Rightarrow y_i * (\vec{w} \cdot \vec{x_i} + b) \geq 1$

$\Rightarrow y_i (\vec{w} \cdot \vec{x_i} + b) - 1 \geq 0$



$w_1 x_{10} + w_2 x_{20} + b = 0$

$w_1 x_{1p} + w_2 x_{2p} + b = 0$

$w_1 (x_{10} - x_{1p}) + w_2 (x_{20} - x_{2p}) = 0$

$\Rightarrow \vec{w} \cdot (\vec{x_0} - \vec{x_p}) = 0$

$y_i (\vec{w} \cdot \vec{x_i} + b) - 1 \geq 0$

$y_i (\vec{w} \cdot \vec{x_i}) \geq \begin{matrix} x_m, & y_i = 1 \\ x_n, & y_i = -1 \end{matrix} \begin{cases} 1+b \\ 1-b \end{cases}$

$$L = \frac{2}{|\vec{w}|} = \frac{|\overset{1+b}{\vec{x_m}} - \overset{1-b}{\vec{x_n}}| \cdot |\vec{w}|}{|\vec{w}|} = \frac{|(1+b) - (1-b)| \cdot |\vec{w}|}{|\vec{w}|} = \frac{|2b| \cdot |\vec{w}|}{|\vec{w}|}$$

$$\text{Max } L = \frac{2}{|\vec{w}|} \leftarrow \text{Min}$$

$$\text{Min } \frac{1}{2} |\vec{w}|^2$$ 推導方便

$$f_i(w, b) = y_i * (\vec{w} \cdot \vec{x_i} + b) - 1 = 0$$ $1 \sim S$

$$L = L(w, b, \alpha_i) = \frac{|\vec{w}|^2}{2} - \sum_{i=1}^{S} \alpha_i [\, y_i (\vec{w} \cdot \vec{x_i} + b) - 1)$$

$$\frac{\partial L}{\partial \vec{w}} = \vec{w} - \sum_{i=1}^{S} \alpha_i y_i x_i = 0 \Rightarrow \vec{w} = \sum_{i=1}^{S} \alpha_i y_i x_i$$

$$\frac{\partial L}{\partial b} = -\sum \alpha_i y_i \Rightarrow \sum \alpha_i y_i = 0$$

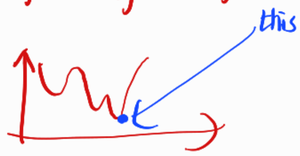
$$L = \frac{1}{2} \left(\sum_{i=1}^{S} \alpha_i y_i \vec{x_i} \right) \left(\sum_{j=1}^{S} \alpha_j y_j \vec{x_j} \right) - \left(\sum_{i=1}^{S} \alpha_i y_i x_i \right) \cdot \left(\sum_{j=1}^{S} \alpha_j y_j x_j \right) - \underbrace{\sum \alpha_i y_i}_{0} b + \sum_{i=1}^{S} \alpha_i$$

$$= \sum \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j \vec{x_i} \cdot \vec{x_j}$$

MLP

$\frac{\partial L}{\partial w_{ij}} = 0$

$w_{ij}^{*} = w_{ij} + \Delta w_{ij}$

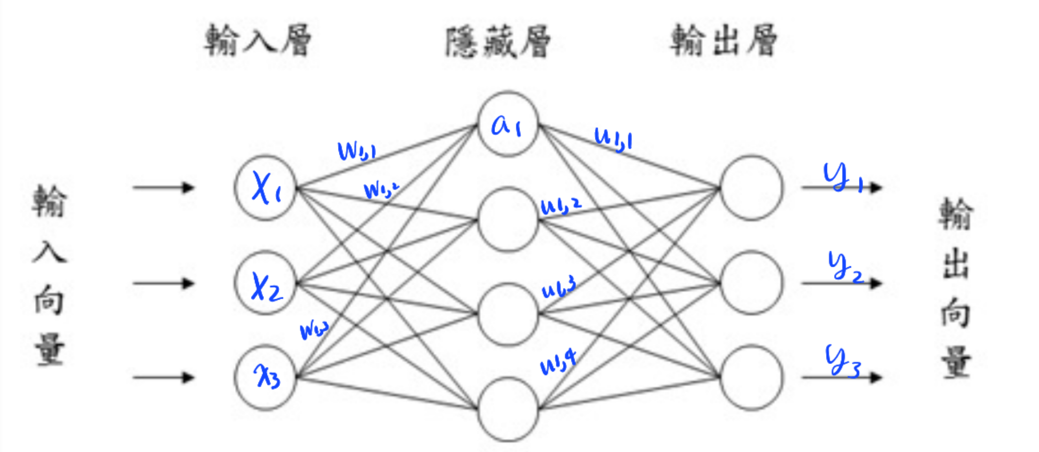


Stochastic Gradient

Rando

SGD

隨機抓樣本



$y_1 = f_2(\vec{a} \cdot \vec{u_1} + b_{21}) \Rightarrow f_2(f_1(\vec{x} \cdot \vec{w_1} + b_{11}) \cdot \vec{u_1} + b_{21})$

$\vec{a} = f_1(\vec{x} \cdot \vec{w_1} + b_{11})$

$$\frac{\partial y_1}{\partial \vec{w_1}} = \frac{\partial y_1}{\partial \vec{a}} \cdot \frac{\partial \vec{a}}{\partial \vec{w_1}}$$

對結果 softmax 正規化 $\sim N(0,1)$ ex. $y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_k \end{bmatrix}$ $S(y_i) = \frac{e^{y_i}}{\sum_{j=1}^{k} e^{y_j}}$ $P = \begin{bmatrix} S(y_1) \\ S(y_2) \\ S(y_3) \\ \vdots \\ S(y_k) \end{bmatrix}$

ArgMAX (P) = 預測結果 $= y$

情況1

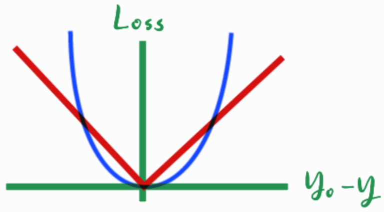
P		y	y_0
0.3		0	0
0.5	$\Rightarrow$	1	1
0.2		0	0

情況2

P		y	y_0
0.3		0	1
0.6	$\Rightarrow$	1	0
0.1		0	0

$\Rightarrow$ y的權重不一樣 所以根據正確答案取 Loss function

損失函數 E 越小越好
- 絕對值損失函數 $\sum_{i=1}^{k} |y_0 - y|$ 損失值隨誤差等比增加
- 平方損失函數 $\sum_{i=1}^{k} (y_0 - y)^2$ 更厭惡大誤差
- $\vdots$



KL散度 $\log y_0 - \log y = \log\left(\frac{y_0}{y}\right)$

ex $E^2 = (\vec{y_0} - \vec{y})^2$

$$\frac{\partial E^2}{\partial \vec{w_1}} = \frac{\partial E^2}{\partial y_1} \cdot \frac{\partial y_1}{\partial \vec{a}} \cdot \frac{\partial \vec{a}}{\partial \vec{w_1}}$$

$y_1 = f_2(\vec{a} \cdot \vec{u_1} + b_{21}) \Rightarrow f_2(f_1(\vec{x} \cdot \vec{w_1} + b_{11}) \cdot \vec{u_1} + b_{21})$