HW3

Name: Yanan Liu

NetID: YL2248

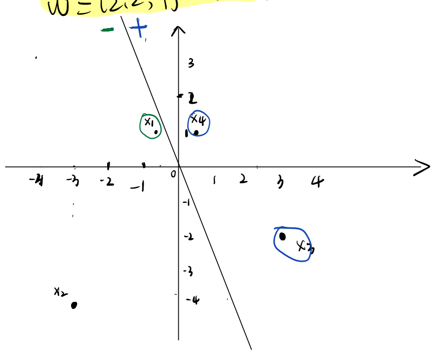
Problem 1:

1.1

start $\vec{w}$	t	x_i	y_i	predicted labels	new $\vec{w}$	new t
(0,1)	0	(-0.6, 1)	-1	+1	(0.6, 0)	1
(0.6, 0)	1	(-3, -4)	-1	-1	(0.6, 0)	1
(0.6, 0)	1	(3, -2)	+1	+1	(0.6, 0)	1
(0.6, 0)	1	(0.5, 1)	+1	-1	(1.1, 1)	0
(1.1, 1)	0	(-0.6, 1)	-1	+1	(1.7, 0)	1
(1.7, 0)	1	(-3, -4)	-1	-1	(1.7, 0)	1
(1.7, 0)	1	(3, -2)	+1	+1	(1.7, 0)	1
(1.7, 0)	1	(0.5, 1)	+1	-1	(2.2, 1)	0
(2.2, 1)	0	(-0.6, 1)	-1	-1	(2.2, 1)	0
(2.2, 1)	0	(-3, -4)	-1	-1	(2.2, 1)	0
(2.2, 1)	0	(3, -2)	+1	+1	(2.2, 1)	0
(2.2, 1)	0	(0.5, 1)	+1	+1	(2.2, 1)	0

Therefore when

$\vec{w} = (2.2, 1)$ $t = 0$, no mistake



Problem 2.

2.1

Let assume $W = ax+by$, if $(a^2+b^2=1)$

this equation means a unit circle. which means

the $r=1$.

For $W_{opt}\cdot X_i - t_{opt} = 0$ $\quad |W_{opt}|^2 = 1$

therefor $W_{opt} X_i = t_{opt}$ $\quad |W_{opt}| = \frac{|t_{opt}|}{X_i} \le 1$

as $X_i \le 1$, $\therefore |t_{opt}| \le 1$

therefore. $|\vec{W_{opt}}|^2 + |\vec{t_{opt}}|^2 \le 2$

Problem 4:

4.1

For any $i = 1, 2, \cdots n$;

$$P(y_1 \cdots y_n | \vec{x_i}, \vec{w}) = \prod_{i=1}^{n} P(y_i | \vec{x_i}, \vec{w})$$

$$= \prod_{i=1}^{n} \frac{1}{2} \left(1 + \frac{y_i(\vec{w}\cdot\vec{x_i})}{\sqrt{1+(\vec{w}\cdot\vec{x})^2}}\right)$$

by taking the log function on both sides.

$$= \sum_{i=1}^{n} \left[\log\frac{1}{2} + (1 + \frac{y_i(\vec{w}\cdot\vec{x})}{\sqrt{1+(\vec{w}\cdot\vec{x})^2}}\right]$$

$$= \sum_{i=1}^{n} \log\frac{1}{2} + \sum_{i=1}^{n} \log\left(1 + \frac{y_i(\vec{w}\cdot\vec{x_i})}{\sqrt{1+(\vec{w}\cdot\vec{x})^2}}\right)$$

$$= \underbrace{n\log\frac{1}{2}}_{-n\log 2} + \sum_{i=1}^{n} \log\left(1 + \frac{y_i(\vec{w}\cdot x_i)}{\sqrt{1+(\vec{w}\cdot x_i)^2}}\right)$$

therefore

$$J(\vec{w}) = -n\log_2 + \sum_{i=1}^{n} \log\left(1 + \frac{y_i(\vec{w}\cdot x_i)}{\sqrt{1+(\vec{w}\cdot x_i)^2}}\right)$$

4.2.

$$\nabla J(\vec{w}) = \frac{d}{d\vec{w}}\left[-n\log_2 + \sum_{i=1}^{n}\log\left(1+\frac{y_i(\vec{w}\cdot x_i)}{\sqrt{1+(\vec{w}\cdot x_i)^2}}\right)\right]$$

$$= \frac{d}{d\vec{w}}\left[\sum_{i=1}^{n}\log\left(1+\frac{y_i(\vec{w}\cdot x_i)}{\sqrt{1+(\vec{w}\cdot x_i)^2}}\right)\right]$$

$$= \sum_{i=1}^{n} \frac{1}{1+\frac{y_i(\vec{w}\cdot\vec{x_i})}{\sqrt{1+(\vec{w}\cdot\vec{x_i})^2}}} \times \frac{(\sqrt{1+(\vec{w}\cdot\vec{x_i})^2})\cdot y_i(\vec{x_i}) - \frac{y_i(\vec{w}\cdot\vec{x_i})(\vec{x_i}\vec{w}\vec{x_i})^2}{\sqrt{1+(\vec{w}\cdot\vec{x_i})^2}}}{(1+(\vec{w}\cdot\vec{x_i})^2)}$$

$$= \sum_{i=1}^{n} \frac{1}{1+\frac{y_i(x_i\cdot\vec{w})}{\sqrt{1+(\vec{w}\cdot x_i)^2}}} \cdot \frac{y_i x_i\sqrt{1+(\vec{w}\cdot x_i)} - y\frac{(\vec{w}\cdot x_i)^2}{\sqrt{1+(\vec{w}\cdot x_i)^2}}}{1+(\vec{w}\cdot x_i)^2}$$

$$= \sum_{i=1}^{n} \frac{1}{1+\frac{y_i(x_i\cdot\vec{w})}{\sqrt{1+(\vec{w}\cdot x_i)^2}}} \cdot \frac{y_i(\vec{x_i})\left(\sqrt{1+(\vec{w}\cdot\vec{x_i})^2} - \frac{(\vec{w}\cdot x_i)^2}{\sqrt{1+(\vec{w}\cdot x_i)^2}}\right)}{1+(\vec{w}\cdot x_i)^2}$$

$$= \sum_{i=1}^{n} \frac{1}{1+\frac{y_i(x_i\cdot\vec{w})}{\sqrt{1+(\vec{w}\cdot x_i)^2}}} \cdot \frac{y_i(\vec{x_i})\cdot\frac{1+(\vec{w}\cdot x_i)^2-(\vec{w}\cdot x_i)^2}{\sqrt{1+(\vec{w}\cdot x_i)^2}}}{1+(\vec{w}\cdot\vec{x_i})^2}$$

$$= \sum_{i=1}^{n} \frac{1}{1+\frac{y_i(\vec{w}\cdot x_i)}{\sqrt{1+(\vec{w}\cdot x_i)^2}}} \cdot \frac{y_i(\vec{x_i})}{(1+(\vec{w}\cdot\vec{x_i})^2)\sqrt{1+(\vec{w}\cdot\vec{x_i})^2}}$$

$$= \sum_{i=1}^{n} \frac{\sqrt{1+(\vec{w}\cdot x_i)^2}}{\sqrt{1+(\vec{w}\cdot x_i)^2}+y_i(x_i\cdot\vec{w})} \cdot \frac{y_i(\vec{x_i})}{(1+(\vec{w}\cdot\vec{x_i})^2)\sqrt{1+(\vec{w}\cdot\vec{x_i})^2}}$$

$$= \sum_{i=1}^{n} \frac{y_i(x_i)}{\sqrt{(1+(\vec{w}\cdot\vec{x_i})^2)^3} + y_i(\vec{w}\cdot\vec{x_i})(1+(\vec{w}\cdot x_i)^2)}$$

therefore.

$$\boxed{\vec{w}_{j+1} = \vec{w_j} + \eta * \sum_{i=1}^{n} \frac{y_i(x_i)}{[(1+(\vec{w}\cdot x_i)^2]^{\frac{3}{2}} + y_i(\vec{w}\cdot\vec{x_i})(1+(\vec{w}\cdot x_i)^2)}}$$

Problem 3.

3.1

Let $x = e^{\frac{a}{2}}$, $y = e^{\frac{b}{2}}$

$\because x^2+y^2 \ge 2xy$ (For $x>0$ & $y>0$)

$\therefore x^2+y^2 = e^a + e^b \ge 2xy = 2e^{\frac{a+b}{2}}$

therefore $\exp(a) + \exp(b) \ge 2\exp(\frac{a+b}{2})$

2.2.

suppose jth update when (x_j, y_j)

$\therefore \vec{w_j} = \vec{w}_{j-1} + y_j \vec{x_i} \qquad t_j = t_{j-1} - y_i$

$\therefore (\vec{w_j}, t_j) \cdot (\vec{w}_{opt}, t_{opt})$

$= (\vec{w}_{j-1} + y_i \vec{x_i},\ t_{j-1} - y_i) \cdot (\vec{w}_{opt}, t_{opt})$

$= \underbrace{\vec{w}_{j-1} \cdot \vec{w}_{opt}}\ + y_i \vec{x_i}\, \vec{w}_{opt} + \underbrace{t_{j-1} t_{opt}} - y_i t_{opt}$

$=$

$= (\vec{w}_{j-1}, t_{j-1}) \cdot (\vec{w}_{opt}, t_{opt}) + y_i (\vec{x_i} \vec{w}_{opt} - t_{opt})$.

$\geq (\vec{w}_{j-1}, t_{j-1}) \cdot (\vec{w}_{opt}, t_{opt}) + r$; since $y_i (x_i w_{opt} - t_{opt}) \geq r$

when $j = j-1$

$(\vec{w}_{j-2}, t_{j-2}) \cdot (\vec{w}_{opt}, t_{opt}) + 2R \geq (j-1) r$

$\therefore (\vec{w}_{j-1}, t_{j-1}) \cdot (\vec{w}_{opt}, t_{opt}) \geq (j-1) r - 2R + r$

$\therefore (\vec{w}_{j-1}, t_{j-1}) (\vec{w}_{opt}, t_{opt}) \geq jr - 2R$.

therefore $(\vec{w_j}, t_j) \cdot (\vec{w}_{opt}, t_{opt}) \geq jr - 2R$.

2.3.

we know $\vec{w_j} = \vec{w}_{j-1} + y_i \vec{x_i} \qquad t_j = t_{j-1} - y_i$

$(\vec{w_j}, t_j)(\vec{w_j}, t_j)$

$= (\vec{w}_{j-1} + y_i x_i,\ t_{j-1} - y_i)(\vec{w}_{j-1} + y_i x_i,\ t_{j-1} - y_i)$

$= (\vec{w}_{j-1} \cdot \vec{w}_{j-1}) + 2 y_i (\vec{w}_{j-1} \cdot x_i) + y_i^2 (x_i \cdot x_i) + (t_{j-1} \cdot t_{j-1}) - y_i t_{j-1} - y_i t_{j-1} + y_i^2$

$= (\vec{w}_{j-1}, t_{j-1}) \cdot (\vec{w}_{j-1}, t_{j-1}) + \vec{x_i}^2 + 2 y_i \underbrace{(\vec{w}_{j-1} x_i - t_{j-1})}_{\leq 0}$

$\leq (\vec{w}_{j-1}, t_{j-1})(\vec{w}_{j-1}, t_{j-1}) + 2$

$\leq 2j - 2 + 2R^2 + 2 = 2j + 2R^2$.

therefore $(\vec{w_j}, t_j) \cdot (\vec{w_j}, t_j) \leq 2j + 2R^2$.

2.4.

We know $|\cos\theta| < 1$

so $\cos\theta = \dfrac{((\vec{w_j}, t_j) \cdot (\vec{w}_{opt}, t_{opt}))^2 \cdots ①}{|\vec{w_j}, t_j|^2\, |w_{opt}, t_{opt}|^2} \leq 1$

from 2.1 we can get ① $\geq jr - 2R$.

from 2.2. we can get $(\vec{w_j}, t_j) \cdot (\vec{w_j}, t_j) \leq 2j + 2R^2$.

therefore we can get $\dfrac{(jr - 2R)^2}{(2j + 2R)^2 \cdot 2} \leq 1$

$j^2 r^2 - 4jrR + 4R^2 \leq 4j + 4R^2$.

$jr^2 - 4rR \leq 4$

$j \leq \dfrac{4 + 4rR}{r^2}$

therefore, we can get $j \leq \dfrac{4 + 4rR}{r^2}$

Name = Yanan Liu

NetID = YL 2248

3.2.

we know

$e^a + e^b \geq 2e^{\frac{a+b}{2}}$

let $a = \vec{w_1} \vec{x}$; $b = \vec{w_2} \cdot \vec{x_2}$

$\therefore \exp(\vec{w_1} \cdot \vec{x}) + \exp(\vec{w_2} \cdot \vec{x}) \geq 2 \exp\left(\dfrac{\vec{w_1} \vec{x} + \vec{w_2} \vec{x}}{2}\right)$

$\therefore \exp(\vec{w_1}, \vec{x}) + \exp(\vec{w_2}, \vec{x}) \geq 2 \exp\left(\dfrac{(\vec{w_1} + \vec{w_2})}{2} \cdot \vec{x}\right)$

3.3.

$J(\bar{w}, t) = \sum_{i=1}^{n} \log P(y_i | \bar{x_i}, \bar{w}, t)$

$= \sum_{i=1}^{n} \left[\dfrac{(1 + y_i)}{2} \cdot (\bar{w} \bar{x_i} - t) - \log(1 + \exp(\bar{w} x_i t)) \right]$

1) let $i = 1$

$\frac{1}{2}\left(\frac{(1+y_1)}{2} \cdot (\bar{w_1} \cdot \bar{x_1} - t_1)\right) + \frac{1}{2}\left(\frac{1+y_1}{2}\right)(\bar{w_2} \cdot \bar{x_1} - t_2) - \log(1 + \exp(\bar{w_1} x_i - t_2)) \Big) \leq \left(\left(\frac{1+y_1}{2} \cdot \left(\frac{\bar{w_1} + \bar{w_2}}{2} \cdot \bar{x_1} - \frac{t_1 + t_2}{2} \right) - \log(1 + \exp\left(\frac{w_1 + w_2}{2} x_1 - \frac{t_1 + t_2}{2}\right) \right)$

We know $\bar{w}^*$ and $\bar{x}_i^* = (\bar{x_i}, -1)$; $\bar{w} x_i - t = \bar{w}^* \cdot \bar{x}_i^*$

$\therefore \frac{1}{2}\left(\frac{(1+y_1)}{2} \cdot (\bar{w}_1^* \cdot \bar{x}_1^*) - \log(1 + \exp(\bar{w}_2^* \cdot \bar{x}_1^*))\right) + \frac{1}{2}\left(\frac{1+y_1}{2} (\bar{w}_1^* \cdot \bar{x}_1^*) - \log(1 + \exp(\bar{w}_2^* \cdot \bar{x}_1^*)\right) \leq$

$\left(\frac{(1+y_1)}{2} \cdot \left(\frac{1}{2}(\bar{w}_1^* + \bar{w}_2^*)\bar{x}_1^*\right) - \log(1 + \exp(\frac{1}{2}(\bar{w}_1^* + \bar{w}_2^*) \cdot \bar{x}_1^*))\right)$

$\dfrac{1+y_1}{4} \cdot (\bar{w}_1^* \cdot \bar{x}_1^* + \bar{w}_2^* \cdot \bar{x}_1^*) - \frac{1}{2}\log(1 + \exp(\bar{w}_1^* \bar{x}_1^*)) - \frac{1}{2}\log(1 + \exp(\bar{w}_2^* \bar{x}_1^*))$

$\leq \dfrac{(1+y_1)}{4} \left((\bar{w}_1^* + \bar{w}_2^*) \cdot \bar{x}_1^* \right) - \log(1 + \exp(\frac{1}{2}(\bar{w}_1^* + \bar{w}_2^*) \cdot \bar{x}_1^*))$

$-\frac{1}{2}\log(1 + \exp(\bar{w}_1^*, \bar{x}_1^*)) - \frac{1}{2}\log(1 + \exp(\bar{w}_2^*, \bar{x}_1^*)) \leq -\log(1 + \exp(\frac{1}{2}(\bar{w}_1^* + \bar{w}_2^*) \cdot \bar{x}^*))$

$\log((1 + \exp(\bar{w}_1^* \cdot x_1^*))(1 + \exp(\bar{w}_2^* \cdot x_i))) \geq 2\log(1 + \exp(\frac{1}{2}(\bar{w}_1^* + \bar{w}_2^*) \cdot \bar{x}_1^*))$

$1 + \exp(\bar{w}_1^* \cdot \bar{x}_1^*)(1 + \exp \bar{w}_2^* \bar{x}_1^*) \geq (1 + \exp(\frac{1}{2}(\bar{w}_1^* + \bar{w}_2^*)\bar{x}_1^*))^2$

$1 + \exp(\bar{w}_1^* \bar{x}_1^*) + \exp(\bar{w}_2^*, \bar{x}_2^*) + \exp((\bar{w}_1^* + \bar{w}_2^*)\bar{x}_1^*) \geq 1 + 2\exp(\frac{1}{2}(\bar{w}_1^* + \bar{w}_2^*) \cdot x_1^*)$

$\exp(\bar{w}_2^*, x_1^*) + \exp(\bar{w}_1^*, x_1^*) \geq 2 \exp(\frac{1}{2}(\vec{w}_1^* + \vec{w}_2^*) \cdot \bar{x}_1^*)$

From 2.2, we have $\exp(\vec{w_1} \cdot \vec{x}) + \exp(\vec{w_2}, \vec{x}) \geq 2\exp\left(\dfrac{(\vec{w_1} + \vec{w_2})}{2} \cdot \vec{x}\right)$

$\therefore$ for $i = 1$, the function is concave.

as concave function + concave function is also concave function
(from google)

therefore:

$J(\bar{w}, t) = \sum_{i=1}^{n} \left[\dfrac{(1+y_i)}{2} (\bar{w} \bar{x_i} - t) - \log(1 + \exp(\bar{x_i} \bar{w} - t)) \right]$ is

concave.