CS 419 Tutorial

1)

i) $\max\{0, 1 - y_i h(x_i)\}$ ✓

ii) $\min\{0, 1 - y_i h(x_i)\}$ ✗

iii) $\dfrac{\exp(-y_i h(x_i))}{1 + \exp(-y_i h(x_i))}$ ✓

iv) $\dfrac{1}{1 + \exp(-y_i h(x_i))}$ ✗

2) $\min\limits_{\theta} \sum_{i=1}^{N} \max\{0, 1 - y_i(w^T x + b)\}$

$\theta \doteq \{w, b\}$

3) a ✓

c

d ✓

bv ~~a~~

ed

4) +1 -1

90% 10%

$h(x) = +1$

$$\hat{h}(x_i) = y_i \rightarrow \text{Best on train}$$

$$h^*(x_i) = y_i \rightarrow \text{Best on test}$$

Answer = 0

5) Without any prior on the test set, we need $r_i = 9$; $y_i = -1$

$r_i = 1$; $y_i = +1$

if 60% $r_i < 9$; $y_i = -1$

$r_i = 1$; $y_i = +1$

6) a) By cases

b) $h(x) = \dfrac{1}{1+e^{-w^T x}}$

$0 < h(x) < 1$

$\Rightarrow L_{term} \geq 0$

$\Rightarrow \forall i \; L_i = 0$

$\Rightarrow$ if $y_i = 1$, $h(x_i) = 1$; if $y_i = 0$, $h(x_i) = 0$ $\Big] \rightarrow w \rightarrow \infty$

c) No.

$\Rightarrow$ Change loss

function!

7) $h(x_i) > \tau$

$\rightarrow$ Obvious flaws with 1/2

→ Why not 3?

→ Can we do "better"? What does "better" even mean?

$$\rightarrow z^* = \underset{z \in [0,1]}{\operatorname{argmin}} \sum \left[\operatorname{sign}(h(x_i) - c) \neq y_i \right]$$

8) Params

a) w, x

b) w

c) w

d) x

e) w, x

f) w', x'

g) $L_2(w,b) = \sum_{i=1}^{N} (y_i - wx_i - b)^2$

$$\hat{w}^*: \left.\frac{\partial L_2(w,b)}{\partial w}\right|_{\hat{w}^*, \hat{b}} = 0$$

$$\sum_{i=1}^{N} 2(y_i - wx_i - b) \cdot x_i = 0$$

$$\Rightarrow \sum y_i x_i - \hat{w}^* \sum x_i^2 - \sum b x_i = 0$$

$$\Rightarrow \hat{w}^* = \frac{\sum y_i x_i - \sum b x_i}{}$$

$$\Rightarrow w \quad \cdots \quad \frac{\quad}{\sum x_i^2}$$

$$\left.\frac{\partial L_2(w,b)}{\partial b}\right|_{b^*} = 0$$

$$\sum (y_i - w x_i - b) = 0$$

$$\Rightarrow b = \frac{\sum y_i}{N} - w^* \frac{\sum x_i}{N}$$

$$\hat{b} = \bar{y} - \hat{w}\,\bar{x}$$

$$\frac{\partial \sum\left((y-\bar{y})-(x-\bar{x})\omega\right)^2}{\partial \omega} = 0$$

$$\boxed{\frac{\sum(y_i-\bar{y})(x_i-\bar{x})}{\sum(x_i-\bar{x})^2}}$$

10) $L_2(\omega) = \sum_{i=1}^{N}(y_i - \omega^T x_i)^2$

$$\left\| \begin{bmatrix} y_1 - w^T x_1 \\ y_2 - w^T x_2 \\ \vdots \\ y_N - w^T x_N \end{bmatrix} \right\|_2^2 = L_2(w)$$

$$X = \begin{bmatrix} x_1^T \\ \vdots \\ x_n^T \end{bmatrix}$$

$$\| Xw - y \|^2$$

① $$\frac{\partial \|x\|^2}{\partial x} = 2x$$

② $$\frac{\partial (Ax)}{\partial x} = A^T$$

$$\Rightarrow \quad X^T(Xw - Y) = 0$$

$$X^T X w = X Y$$

$$\Rightarrow w = (X^T X)^{-1} X Y$$

1) $X^T X$ is invertible

when X is a full

column rank

$$A^TAx = 0 \Rightarrow x^TA^TAx = 0$$

$$\Rightarrow \|Ax\| = 0$$

$$\Rightarrow Ax = 0$$

$$\Rightarrow x = 0$$

(a) $(X^T X + \lambda I)$

$$v^T(X^TX+\lambda I)v$$

$$v^TX^TXv + \lambda v^Tv$$

$$\underbrace{\|Xv\|}_{\geq 0} + \lambda\underbrace{\|v\|}_{\geq 0}$$

3) $y_i = w^T x_i + \epsilon_i$

$y_i \sim N(w^T x_i, \sigma^2)$

$$P_i = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y_i - w^T x_i)^2}{2\sigma^2}}$$

$(y_i - w^T x_i)^2$

$$\prod P = \frac{1}{\sqrt{2\pi\sigma}^{\,n}} e^{-\frac{\sum(\ldots)}{2\sigma^2}}$$

$$\boxed{\min \sum (y_i - w^T x_i)^2}$$