Viveksinh Solanki
ML – 10441787

Q1. Support Vector Machines:

☆ Assignment-3 :-

Q-1 :- (a) We know that equation of hyperplane for svm is

$$\boxed{W^T x + b = 0} \quad \text{—— (1)}$$

- We can use, $W = \sum_{i=1}^{n} \alpha_i y_i x_i$, to find W.

$$= \alpha_1 y_1 x_1 + \alpha_2 y_2 x_2 + \dots + \alpha_{10} y_{10} x_{10} \quad \text{—— (2)}$$

But, $\alpha_2 = \alpha_3 = \alpha_5 = \alpha_6 = \alpha_8 = \alpha_{10} = 0$

∴ from —— (2)

$$W = \alpha_1 y_1 x_1 + \alpha_4 y_4 x_4 + \alpha_7 y_7 x_7 + \alpha_9 y_9 x_9$$

$$= (0.414)(1)\begin{bmatrix}4\\2.9\end{bmatrix} + (0.018)(-1)\begin{bmatrix}2.5\\1\end{bmatrix} + (0.018)(1)\begin{bmatrix}3.5\\4\end{bmatrix} + (0.414)(-1)\begin{bmatrix}2\\2.1\end{bmatrix}$$

$$= (0.414)\left\{\begin{bmatrix}4\\2.9\end{bmatrix} - \begin{bmatrix}2\\2.1\end{bmatrix}\right\} + (0.018)\left\{\begin{bmatrix}3.5\\4\end{bmatrix} - \begin{bmatrix}2.5\\1\end{bmatrix}\right\}$$

$$= 0.414\begin{bmatrix}2\\0.8\end{bmatrix} + 0.018\begin{bmatrix}1\\3\end{bmatrix}$$

$$= \begin{bmatrix}0.828\\0.33\end{bmatrix} + \begin{bmatrix}0.018\\0.054\end{bmatrix}$$

$$= \begin{bmatrix}0.846\\0.384\end{bmatrix} = \begin{bmatrix}w_1\\w_2\end{bmatrix} \quad \text{—— (3)}$$

Viveksinh Solanki
ML – 10441787

→ To calculate b, we can plug all available values in,

$\alpha_i \left[y_i (w^T x_i + b) - 1 \right] = 0$, for any of the support vectors

→ Using x_1,

$$(0.414)\left[(1)\left\{\begin{bmatrix}0.846\\0.384\end{bmatrix}^T\begin{bmatrix}4\\2.9\end{bmatrix} + b\right\} - 1\right] = 0$$

$$\therefore \cancel{(2.453 + 1.536) + b = 1}$$

$$\cancel{\therefore b = 1 - 3.989}$$

$$\cancel{\therefore b = -2.989}$$

$$1 - \begin{bmatrix}0.846 & 0.384\end{bmatrix}\begin{bmatrix}4\\2.9\end{bmatrix} = b$$

$$\therefore b = 1 - (3.384 + 1.1136)$$

$$= 1 - 4.4976$$

$$= -3.497$$

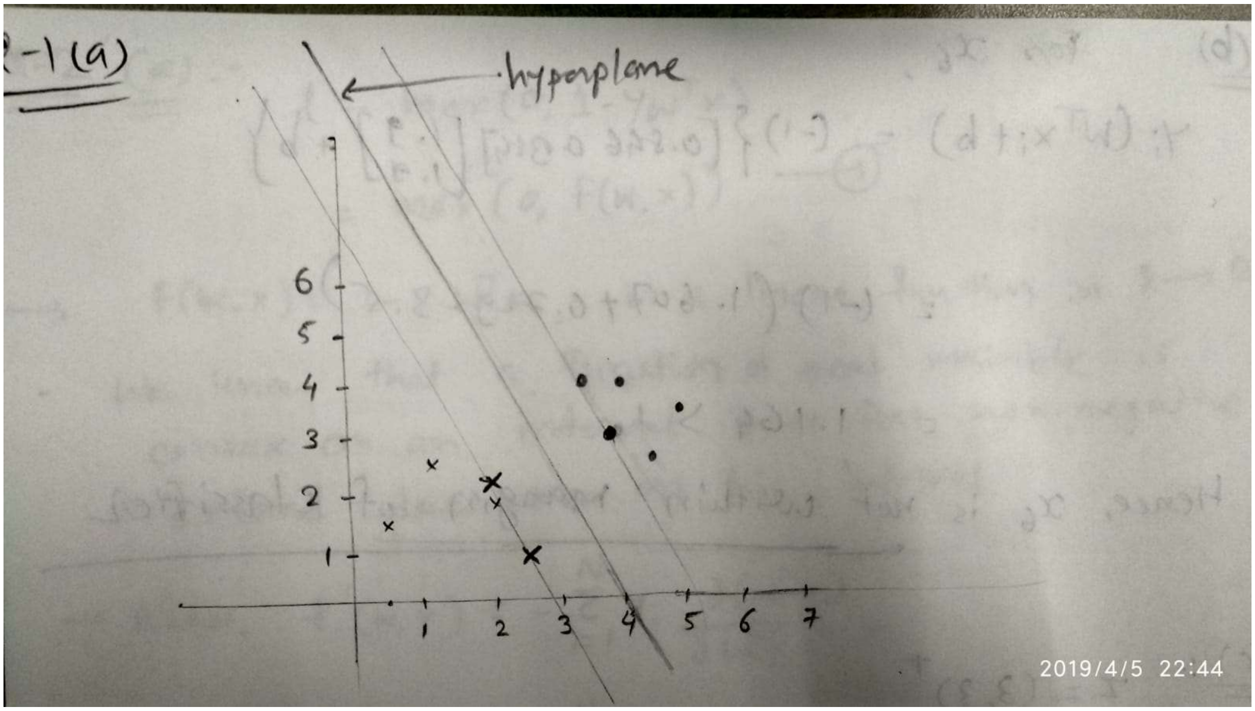
$$\boxed{b \approx -3.5}$$

→ By plugging w & b values into ①,

$$\begin{bmatrix}0.846 & 0.384\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} + (-3.5) = 0$$

$$\therefore \boxed{0.846x_1 + 0.384x_2 - 3.5 = 0}$$ — hyperplane.

Viveksinh Solanki
ML – 10441787



Viveksinh Solanki
ML – 10441787

Q-1(b) We know that, $d(x, \text{hyperplane}) = \frac{h(x)}{\|w\|}$,

$$\Rightarrow h(x) = w^T x + b = 0.846x_1 + 0.384x_2 - 3.5 \quad \text{(from-①(a))}$$

$$\therefore d(x_6, h(x)) = \frac{0.846(1.9) + 0.384(1.9) - 3.5}{\sqrt{0.846^2 + 0.384^2}}$$

$$= \frac{-1.163}{\sqrt{0.715 + 0.147}}$$

$$= \frac{-1.163}{\sqrt{0.862}}$$

$$= \boxed{-1.253},$$

~~Hence it is not fitting margin of the classifier.~~

Viveksinh Solanki
ML – 10441787

Q-1 (b) for x_6,

$\Rightarrow \quad y_i (W^T x_i + b) = (-1) \left\{ [0.846 \;\; 0.384] \begin{bmatrix} 1.9 \\ 1.9 \end{bmatrix} + b \right\}$

$= (-1)(1.607 + 0.729 - 3.5)$

$= 1.164 > 1$

- Hence, x_6 is not within margin of classifier

Q-1 (c) :- $z = (3, 3)^T$

$\rightarrow \quad h(x) = 0.846 x_1 + 0.384 x_2 - 3.5$

$= 0.846(3) + 0.384(3) - 3.5$

$= 2.53 + 1.15 - 3.5$

$= 0.18 > 0$

hence, z is on positive side of hyperplane.

$\therefore$ for given z, $\boxed{y = 1}$

Viveksinh Solanki
ML – 10441787

Q2. Support Vector Machines:

Q-2 (a) :-
$$l = \max(0, 1 - YW^T x)$$
$$= \max(0, f(W, x)) \quad \text{--- (1)}$$

→ $f(W.x) = 1 - y.W^T x$, is a linear function in $R \to R$

- We know that a function of real variable is convex on an interval if it has non-negative second derivative on this interval.

→ Now, $f'(W, x) = -\sum_{i=1}^{N} Y_i \frac{\partial (W^T x_i)}{\partial W}$

$$= -\sum_{i=1}^{N} Y_i x_i \quad \text{--- (2)}$$

→ Taking second derivative w.r.t. W

$$f''(W, x) = \frac{\partial}{\partial W}\left(-\sum_{i=1}^{N} Y_i x_i\right)$$
$$= \boxed{0}, \quad \cancel{\forall f''(W,x): R \to R}$$
$$\forall f(W, x) : R \to R \quad \text{--- (3)}$$

→ hence, ~~f(x, w~~ $f(W, x) = 1 - Y.W^T x$, is a convex function.

→ Also, the maximum of two convex functions is a convex function. ← (4)

→ zero(0) is convex and from –(1), –(3) 4 –(4).

$\boxed{l = \max(0, 1 - YW^T x)}$ is a convex function

Viveksinh Solanki
ML – 10441787

Q:- 2(b) :-

→ $l = \max(0, 1 - y_i f(x_i))$, $f(x_i) = \mathrm{sgn}(w^T x_i)$

→ We know that, $\boxed{l \geq 0}$ —①

→ ~~If~~ $f(x_i) = \mathrm{sgn}(w^T x_i)$ is given, which implies [Correct Predictions are given]

$y_i w \cdot x_i \geq 0$, hence

$\boxed{1 - y_i w x_i \leq 1.}$ —②

→ —② gives,

$\boxed{l = \max(0, 1 - y_i w \cdot x_i) \leq 1.}$ —③

⇒ Hence, from —① and —③

$\boxed{0 \leq \max(0, 1 - y_i f(x_i)) \leq 1}$

Q-2(c) :- $M(w) =$ # of mistakes by w

$l = \max(0, 1 - y_i w^T x_i)$

⇒ Here, misclassification is happening.

hence, $y_i w^T x_i \geq 0$, implies,

$(1 - y_i w^T x_i) \geq 1$, implies,

$\boxed{l = \max(0, 1 - y_i w^T x_i) \geq 1}$ —①

Viveksinh Solanki
ML – 10441787

Q-2(C) :- Continued....

$$\therefore \underline{M(w)} = \sum_{i=1}^{n} 1 = \sum_{n} 1 \leq \sum_{n} l$$

$$\leq \sum_{i=1}^{n} (\max(0, 1-y_i w^T x_i))$$

→ Hence, $\frac{1}{n} M(w) \leq \frac{1}{n} \sum_{i=1}^{n} \max(0, 1-y_i w^T x_i)$