$\boxed{9.1:}$

We know that for any $c \in \mathbb{R}$, if $\begin{cases} c \leq a \\ \text{and} \\ c \geq -a \end{cases}$

we can say that $|c| = \min\limits_{a \geq 0} a$

if we take the emprical risk $|\langle w, x_i \rangle - y_i|$ as $\underline{\underline{a}}$ and an auxiliary variable like Z_i as $\underline{\underline{c}}$, then we can conclude that minimizing the expression $\sum_{i=1}^{m} |\langle w, x_i \rangle - y_i|$ is equal to minimizing $\sum_{i=1}^{m} Z_i$ under the following constraints:

$$\forall i \in \{1, \ldots, m\}: \begin{cases} -(\langle w, x_i \rangle - y_i) \leq s_i \\ \text{and} \\ \langle w, x_i \rangle - y_i \geq s_i \end{cases}$$

$$\Rightarrow \begin{cases} -\langle w, x_i \rangle - s_i \leq -y_i \\ \\ \langle w, x_i \rangle - s_i \leq y_i \end{cases}$$

Assume bellow matrices where:

$$A = \left[\begin{array}{c:c} X & -I_m \\ \hdashline -X & -I_m \end{array}\right]_{2m \times (d+m)}$$

$$\mathcal{V} = \begin{bmatrix} w_1 \\ \vdots \\ w_d \\ z_1 \\ \vdots \\ z_m \end{bmatrix}_{(d+m)\times 1} \qquad b = \begin{bmatrix} y_1 \\ \vdots \\ y_m \\ -y_1 \\ \vdots \\ -y_m \end{bmatrix}$$

$$C = \begin{array}{r} d \text{ times} \left\{ \begin{array}{c} \\ \\ \\ \end{array} \right. \\ \begin{array}{c} m \\ \text{time} \end{array} \left\{ \begin{array}{c} \\ \\ \end{array} \right. \end{array} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{(d+m)\times 1}$$

We wished to minimize $\sum_{i=1}^{m} z_i$ which equal to minimizing $C^T \mathcal{V}$. The $\begin{cases} \langle w, x_i \rangle - s_i \leq y_i \\ -\langle w, x_i \rangle - s_i \leq -y_i \end{cases}$ constraints are also equal to $A\mathcal{V} \leq b$ which turns it to LP. $\blacksquare$

9.6

1) Assume there exists an example set like $\underline{\underline{S}}$ that is shattered by the class of closed balls in R^d.

$S = \{x_1, \dots, x_m\}$ shattered by $B_d = \{B_{v,r} : v \in R^d, r > 0\}$

We want to prove that these examples are also shattered by the class of Half Spaces in R^{d+1} with these input formats: $\phi : R^d \longrightarrow R^{d+1}$

$$\phi(x) = (x, \|x\|^2)$$

So we want to prove that $y_i(\langle W(x, \|x\|^2) \rangle + b)$ will be always positive, which means it always classifies the m examples right. In other words:

$$y_i(\langle W(x, \|x\|^2) \rangle + b) \geq 0$$

for $y_i = 0$, the above statement holds. We only need to prove the above state where $y_i = 1$.

Because we know that B_d shatters S,

for $y_i = 1$ we have:

$$\|x - v\| \leq r$$

$$\|x - v\| \leq r^2$$

$$(x - v)^T (x - v) \leq r^2$$

$$\|x\| - v^T x + \|v\| \underbrace{- x^T v}_{\text{equal to } -v^T x} \leq r^2$$

$$\Rightarrow \|x\| - 2 v^T x + \|v\| \leq r^2$$

$$\cdot \leq r^2 - \|x\| + 2 v^T x - \|v\|$$

if you take $W = \left[2v^T \,\vdots\, -1 \right]$ & $b = r^2 - \|v\|$

We just proved that for $y_i = 1$, we have

$$y_i \left(\langle W (x, \|x\|) \rangle + b \right) > 0 \quad \blacksquare$$

2)

Assume $m = d+1$, So as we showed,

If B_d shatteres $m=d+1$ examples $\Rightarrow$

$VCdim(B_d) \geq d+1$

Also we know that these $d+1$ points can be shattered by halfspaces in R^{d+1}.

in other words, $VCdim(B_d) \leq VCdim(HS)$

From Our prior knowledge, we know that

$VCdim(\text{Half Spaces in } R^{d+1}) \leq d+2$

$$\Rightarrow \quad d+1 \leq VCdim(B_d) \leq d+2$$