1, Problem 1.3

a) Let $\rho = \min_{1 \leq n \leq N} y_n (w^{*T} x_n)$. Show that $\rho > 0$

Since w^* separates the data perfectly

↳ $y_n (w^{*T} x_n)$ always $> 0 \Rightarrow \rho > 0$

b) • $w^T(t) w^* \geq w^T(t-1) w^* + \rho$

We have : $w(t) = w(t-1) + y(t-1) x(t-1)$

$\rightarrow w^T(t) = \left(w(t-1) + y(t-1) x(t-1) \right)^T$

$\rightarrow w^T(t) w^* = \left(w(t-1) + y(t-1) x(t-1) \right)^T w^*$

$\rightarrow w^T(t) w^* = w^T(t-1) w^* + \underbrace{y(t-1) x(t-1) w^{*T}}_{\geq \rho}$

$\Rightarrow w^T(t) w^* \geq w^T(t-1) w^* + \rho$

- $w^T(t)w^* \geq t\rho \quad (1)$

  - Base Case: $t = 0$

    ↳ $w^T(0)w^* \geq 0\rho = 0$

    We assume $w(0) = 0$

    $\rightarrow 0 \geq 0$: true

- Induction: Assume (1) is true at t, need to prove (1) is true at $t+1$

$$w(t+1) = w(t) + y(t)x(t)$$

$$\rightarrow w^T(t+1)w^* = \left(w(t) + y(t)x(t)\right)^T w^*$$

$$= \underbrace{w^T(t)w^*}_{\geq t\rho \, (1)} + \underbrace{w^{*T}\left(y(t)x(t)\right)}_{\geq \rho}$$

$$\Rightarrow w^T(t+1)w^* \geq t\rho + \rho = \rho(t+1)$$

↳ true at $t+1$

$$\therefore w^T(t)w^* \geq t\rho$$

c) $\|w(t)\|^2 \leqslant \|w(t-1)\|^2 + \|x(t-1)\|^2$

We have : $\|w(t)\|^2 = \|w(t-1) + y(t-1)x(t-1)\|^2$

$= \left(w(t-1) + y(t-1)x(t-1)\right)\left(w(t-1) + y(t-1)x(t-1)\right)^T$

$= \|w(t-1)\|^2 + \underbrace{(y(t-1))^2}_{1}\|x(t-1)\|^2 + \underbrace{2y(t-1)x(t-1)w^T(t-1)}_{\leqslant 0}$

$\Rightarrow \|w(t)\|^2 \leqslant \|w(t-1)\|^2 + \|x(t-1)\|^2$

d) (2) $\|w(t)\|^2 \leqslant tR^2$, $R = \max_{1 \leqslant n \leqslant N} \|x_n\|$

- Base case : $t = 0$

  $\hookrightarrow \|w(0)\|^2 = 0 = 0R^2$ (assume $w(0) = 0$)

- Induction : Assume (2) is true at t, need to prove (2) is true at $t+1$

$$\left.\begin{array}{l} (t+1)R^2 = tR^2 + R^2 \geqslant \|w(t)\|^2 + R^2 \\ R \geqslant \|x_n\| \rightarrow R^2 \geqslant \|x_n\|^2 \end{array}\right\} \rightarrow$$

$\Rightarrow (t+1)R^2 \geqslant \|w(t)\|^2 + \|x(t)\|^2 \geqslant \|w(t+1)\|^2$

$\hookrightarrow \|w(t+1)\|^2 \leqslant (t+1)R^2 \rightarrow$ (2) is true at $t+1$

(e) $\bullet \dfrac{w^T(t)}{\|w(t)\|} w^* \geq \sqrt{t} \cdot \dfrac{\rho}{R}$

From b, we have : $w^T(t)\, w^* \geq w^T(t-1) w^* + \rho$

$$\hookrightarrow \frac{w^T(t)\, w^*}{\|w(t)\|} \geq \frac{w^T(t-1) w^* + \rho}{\|w(t)\|}$$

$$\geq \frac{(t-1)\rho + \rho}{\|w(t)\|} = \frac{t\rho}{\|w(t)\|} \quad \text{(A)} \; (w^T(t)\, w^* \geq t\rho)$$

From d, we have $\|w(t)\|^2 \leq tR^2 \rightarrow \dfrac{1}{\|w(t)\|^2} \geq \dfrac{1}{tR^2}$

$$\rightarrow \frac{1}{\|w(t)\|} \geq \frac{1}{\sqrt{t}\, R} \quad \text{(B)}$$

$$\text{(A)}, \text{(B)} \rightarrow \frac{w^T(t)\, w^*}{\|w(t)\|} \geq \frac{t\rho}{\sqrt{t}\, R} = \sqrt{t}\, \frac{\rho}{R}$$

- $t \leqslant \dfrac{R^2 \|w^*\|^2}{\rho^2}$

Assume θ is the angle between $w^T(t)$ and w^*

$\hookrightarrow \cos\theta = \dfrac{w^T(t)\, w^*}{\|w^T(t)\|\ \|w^*\|} \leqslant 1$

$$\Rightarrow \quad \|w^*\| \geqslant \frac{w^T(t) w^*}{\|w^T(t)\|} \geqslant \sqrt{t}\, \frac{\rho}{R}$$

$$\Rightarrow \quad \frac{\|w^*\|^2 R^2}{\rho^2} \geqslant t$$

2,

I want to take this class because I am double majoring in CS and Financial Engineering which requires Machine Learning. I am planning on doing a co-op internship next Fall too so if I can take the class this semester, I can apply to more intern positions.