4.

Let PLA finds answer after t+1 times

$\vec{W}_{t+1} \leftarrow \vec{W}_t + y_{n(t)} \vec{X}_{n(t)}$

$\Rightarrow \vec{W}_{t+1} \leftarrow y_{n(t)} \vec{X}_{n(t)} + y_{n(t-1)} \vec{X}_{n(t-1)} + \cdots + y_{n(0)} \vec{X}_{n(0)} + \vec{W}_0$

↳ say $\vec{0}$, from page 8/22 of lecture 2

Since the zero-th component of each $\vec{X}$ is 1,

We get $W_0 = \sum_{k=0}^{t} y_{n(k)}$

then $W_0 = T_+ - T_-$

5.

$T \leq \left(\frac{R}{\rho}\right)^2$ where $R^2 = \max_n \|\vec{X}_n\|^2$, $\rho = \min_n y_n \frac{\vec{W}_f^T}{\|\vec{W}_f\|} \vec{X}_n$

in this situation, $R^2 = \max_n \|\vec{X}_n\|^2 = \sqrt{(m+1)}^2 = m+1$,

$\rho^2 = \frac{1}{\sqrt{d+(0.5)^2}^2}\left(\min_n y_n \vec{W}_f^T \vec{X}_n\right)^2 = \frac{1}{d+1/4}\left(\min_n \operatorname{sign}\left(Z_+(\vec{X}_n) - Z_-(\vec{X}_n) - 0.5\right)\left(Z_+(\vec{X}_n) - Z_-(\vec{X}_n) - 0.5\right)\right)^2$

$= \frac{4}{4d+1}\left(\min_n \left| Z_+(\vec{X}_n) - Z_-(\vec{X}_n) - 0.5 \right|\right)^2$

$\overset{\text{when } Z_+(x) = Z_-(x)}{=} \frac{4}{4d+1}(0.5)^2 = \frac{1}{4d+1}$

therefore $T \leq \frac{R^2}{\rho^2} = (m+1)(4d+1)$ ※

6.

Suppose we have two points, $X_1 = (1, -1)$ and $X_2 = (-1, 1)$ where $y_1 = 1$, $y_2 = -1$

In the first case, $\vec{X}_1 = \langle 1, 1, -1 \rangle$, $\vec{X}_2 = \langle 1, -1, 1 \rangle$

PLA process: $\vec{W}_1 \leftarrow \vec{W}_0 + \vec{X}_1 = \langle 1, 1, -1 \rangle = W_{PLA}$

then we have $\vec{W}_{pla} = \vec{W}_1 = \langle 1, 1, -1 \rangle$

In the second case, $\vec{X}_1 = \langle -1, 1, -1 \rangle$, $\vec{X}_2 = \langle -1, -1, 1 \rangle$

PLA process: $\vec{W}_1 \leftarrow \vec{W}_0 + \vec{X}_1 = \langle -1, 1, -1 \rangle = W_{PLA}'$

therefore we can see $W_{PLA} \neq W_{PLA}'$, we disprove that in case 1 & case 2 the W_{PLA} would be the same.

7.

$T \le R^2/\rho^2$, $R = \max_n \|\vec{z}_n\| = 1$

$\rho = \min_n y_n \frac{W_f^T}{\|W_f\|} z_n$

therefore $T \le \frac{1}{\rho^2}$

8. prove PAM halts in finite steps $\Leftrightarrow$ exist perfect W_f such that for all $y_n = \text{sign}(W_f^T x_n)$

① W_f perfect, hence every x_n correctly away from line:

$$y_n W_f^T x_{n(t)} \ge \min_n y_n W_f^T x_n > 0$$

② $W_T^f W_t \uparrow$ by updating with any $(x_{n(t)}, y_{n(t)})$;

$$\begin{aligned} W_f^T W_{t+1} &= W_f^T (W_t + y_{n(t)} x_{n(t)}) \\ &\ge W_f^T W_t + \min_n y_n W_f^T x_n \\ &> W_f^T W_t + T > W_f^T W_t \quad (T > 0) \end{aligned}$$

③ $\|W_t\|$ doesn't grow too fast ;

$$\begin{aligned} \|W_{t+1}\|^2 &= \|W_t + y_{n(t)} x_{n(t)}\|^2 \\ &= \|W_t\|^2 + 2 y_{n(t)} W_t x_{n(t)} + \|y_{n(t)} x_{n(t)}\|^2 \\ &\le \|W_t\|^2 + 2T + \|y_{n(t)} x_{n(t)}\|^2 \\ &\le \|W_t\|^2 + \underbrace{\max_n \|y_n x_n\|^2 + 2T}_{\text{constant}} \end{aligned}$$

Therefore, PAM halts in finite steps ※