7. Let $\vec{x_n}^{orig} = [x_{n1}\ x_{n2}\ \cdots\ x_{nd}] \in \mathbb{R}^d$ for $n \in \{1, \dots, N\}$

then $\vec{x_n} = [1\ x_{n1}\ x_{n2}\ \cdots\ x_{nd}] \in \mathbb{R}^{d+1}$ for $n \in \{1, \dots, N\}$

and $\vec{x_n}' = [2\ x_{n1}\ x_{n2}\ \dots\ x_{nd}] \in \mathbb{R}^{d+1}$ for $n \in \{1, \dots, N\}$

Suppose for the the $\vec{x_n}$'s that are formed conventionally,

the first p-1 inputs are predicted correctly, then:

$$\text{sign}(\vec{w_0}^T \vec{x_i}) = y_i \quad \text{for } i = 0 \sim p-1$$

If we assume $\text{sign}(0) = -1$ for simplicity, and $\vec{w_0} = \vec{0}$, then this means that.

$$\text{sign}(\vec{0}^T \vec{x_i}) = \text{sign}(0) = -1 = y_i \quad \text{for } i = 1 \sim p-1$$

If we further assume $\vec{x_p}$ receives incorrect prediction, then

$$\text{sign}(\vec{0}^T \vec{x_p}) \neq y_p$$

$$\Rightarrow \text{sign}(0) = -1 \neq y_p \qquad \therefore y_p = 1$$

Thus we update $\vec{w_1}$ as:

$$\vec{w_1} = \vec{w_0} + y_p \vec{x_p} = \vec{0} + 1 \cdot \begin{bmatrix} 1 \\ x_{p1} \\ x_{p2} \\ \vdots \\ x_{pd} \end{bmatrix} = \begin{bmatrix} 1 \\ x_{p1} \\ x_{p2} \\ \vdots \\ x_{pd} \end{bmatrix}$$

Since $\vec{w_0}' = \vec{0}$, and y_i remain unchanged no matter we use $\vec{x_i}$ or $\vec{x_i}'$, so:

$$\text{sign}(\vec{w_0}'^T \vec{x_i}') = \text{sign}(\vec{0}^T \vec{x_i}') = -1 = y_i \quad \text{for } i = 1, \dots, p-1$$

We next want to know if $\vec{x_p}'$ will get correct prediction or not:

$$\text{sign}(\vec{w_0}'^T \vec{x_p}') = \text{sign}(\vec{0}^T \vec{x_p}') = -1 \neq 1 = y_p$$

Thus $\vec{x_p}'$ is also mispredicted, and we update the weight vector as:

$$\vec{w_1}' = \vec{w_0}' + y_p \vec{x_p}' = \vec{0} + 1 \cdot \begin{bmatrix} 2 \\ x_{p1} \\ x_{p2} \\ \vdots \\ x_{pd} \end{bmatrix} = \begin{bmatrix} 2 \\ x_{p1} \\ x_{p2} \\ \vdots \\ x_{pd} \end{bmatrix}$$

Consider the next input $\vec{x_{p+1}}$ and $\vec{x_{p+1}}'$:

$$\boxed{\text{sign}(\vec{w_1}^T \vec{x_{p+1}})} = \text{sign}\left(\begin{bmatrix} 1 & x_{p1} & x_{p2} & \cdots & x_{pd} \end{bmatrix} \begin{bmatrix} 1 \\ x_{p+1,1} \\ x_{p+1,2} \\ \vdots \\ x_{p+1,d} \end{bmatrix} \right)$$

$$= \underline{\text{sign}\left(1 + \sum_{i=1}^{d} x_{pi} x_{p+1,i}\right)}$$

$$\boxed{\text{sign}(\vec{w_1}'^T \vec{x_{p+1}}')} = \text{sign}\left(\begin{bmatrix} 2 & x_{p1} & x_{p2} & \cdots & x_{pd} \end{bmatrix} \begin{bmatrix} 2 \\ x_{p+1,1} \\ x_{p+1,2} \\ \vdots \\ x_{p+1,d} \end{bmatrix} \right)$$

$$= \underline{\text{sign}\left(4 + \sum_{i=1}^{d} x_{pi} x_{p+1,i}\right)}$$

Therefore, consider the situation that if $\sum_{i=1}^{d} x_{pi} x_{p+1,i} = -3$

then:

$$\text{sign}(\vec{w_1}^T \vec{x_{p+1}}) = \text{sign}(1+(-3)) = \text{sign}(-2) = -1$$

$$\text{sign}(\vec{w_1}'^T \vec{x_{p+1}}') = \text{sign}(4+(-3)) = \text{sign}(1) = 1$$

$\Rightarrow$ This will cause $\vec{w_1}$ to update if $y_{p+1} = 1$, $\vec{w_1}'$ to update if $y_{p+1} = -1$.

If the case is $y_{p+1} = -1$, and there are only $p+1$ point in the dataset,

(i.e. $\{\vec{x_n}, y_n\}_{n=1}^{p+1}$ / $\{\vec{x_n}', y_n\}_{n=1}^{p+1}$)

then $\vec{w_{PLA}} = \vec{w_1}$

and we continue to update $\vec{w_1}'$ and check for $\vec{x_1}'$:

$$\vec{w_2}' = \vec{w_1}' + y_{p+1} \cdot \vec{x_{p+1}}' = \begin{bmatrix} 2 \\ x_{p1} \\ x_{p2} \\ \vdots \\ x_{pd} \end{bmatrix} + (-1) \cdot \begin{bmatrix} 2 \\ x_{p+1,1} \\ x_{p+1,2} \\ \vdots \\ x_{p+1,d} \end{bmatrix} = \begin{bmatrix} 0 \\ x_{p1} - x_{p+1,1} \\ x_{p2} - x_{p+1,2} \\ \vdots \\ x_{pd} - x_{p+1,d} \end{bmatrix}$$

$$\text{sign}(\vec{w_2}'^T \vec{x_1}') = \text{sign}\left(\begin{bmatrix} 0 & x_{p1} - x_{p+1,1} & \cdots & x_{pd} - x_{p+1,d} \end{bmatrix} \begin{bmatrix} 2 \\ x_{11} \\ x_{12} \\ \vdots \\ x_{1d} \end{bmatrix} \right)$$

$$= \text{sign}\left(\sum_{i=1}^{d} (x_{pi} - x_{p+1,i}) x_{1i} \right)$$

If $\sum_{i=1}^{d} (x_{pi} - x_{p+1,i}) x_{1i} \leq 0$, and since $y_1 = -1$ as we assumed earlier, we have correct prediction using $\vec{w_2}'$.

We can imagine that we construct the inputs to make:

$$\text{sign}(\vec{w_2}'^T \vec{x_i}) = y_i \quad \text{for } i = 1, \dots, p+1$$

then $\vec{w}_{PLA}' = \vec{w_2}' = \begin{bmatrix} 0 \\ x_{p1} - x_{p+1,1} \\ \vdots \\ x_{pd} - x_{p+1,d} \end{bmatrix}$

or we can see it is trivial that each time we update $\vec{w_i}'$, we add / substract 2 to w_{i0}, thus

$$\vec{w}_{PLA}' = \begin{bmatrix} 2k \\ \vdots \\ \vdots \end{bmatrix} \quad \text{for } k \in \mathbb{Z}$$

and $\vec{w}_{PLA} = \begin{bmatrix} 1 \\ \vdots \\ \vdots \end{bmatrix}$, so $\vec{w}_{PLA}' \neq \vec{w}_{PLA}$. $\square$