ECE 421 Homework 1

Name: Mark Qi

Student ID: 1006764645

Question 1

$D = \{(x_n, y_n)\}_{n=1}^{N}$, $x_n \in \mathbb{R}^d$, $y_n \in \{+1, -1\}$

$$h(x) = \text{sign}\left(b + \sum_{i=1}^{d} w_i x_i\right) = \text{sign}(w^T x)$$

Perceptron weight update rule (1.3)

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

(a)

Since $x(t)$ is misclassified by $w(t)$, this means $h(t) \neq y(t)$, thus one is positive and the other is negative since $h(t) \in \{+1, -1\}$ and $y(t) \in \{+1, -1\}$

Either $h(x) = -1$, $y(x) = 1$

$$y(t)h(x) = y(t)w(t)^T x(t) < 0$$

Or $h(x) = 1$, $y(x) = -1$

$$y(t)h(x) = y(t)w(t)^T x(t) < 0$$

Thus, for all misclassified $x_i(t)$,

$$y_i(t)w(t)^T x_i(t) < 0$$

(b)

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

$$y(t)w(t+1)^T x(t) = y(t)((w(t)+y(t)x(t))^T x(t))$$
$$= y(t)(w(t)^T x(t) + y(t)x(t)^T x(t))$$
$$= y(t)(w(t)^T x(t) + y(t)\|x(t)\|_2^2)$$
$$= y(t)w(t)^T x(t) + y(t)^2 \|x(t)\|_2^2$$
$$y(t) \in \{+1, -1\},\ y(t)^2 = 1$$
$$= y(t)w(t)^T x(t) + \|x(t)\|_2^2$$
$$> y(t)w(t)^T x(t)$$

(c)

Perceptron Weight Update Rule

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

Since $y_n(w_k^T x_n) < y_n(w_{k+1}^T x_n)$, we are moving towards the more positive position

Question 2,

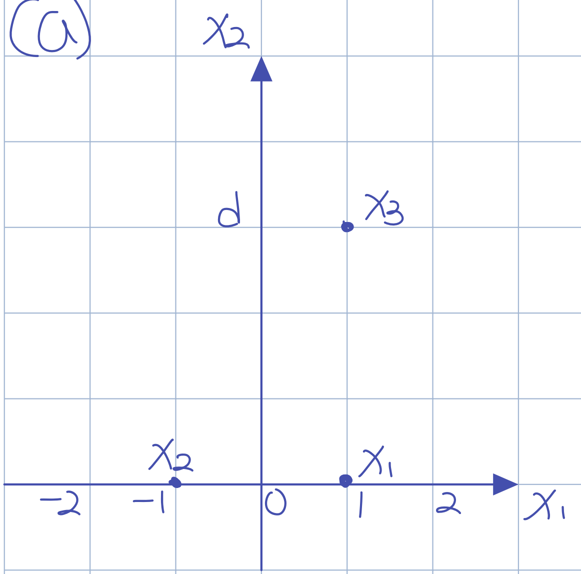
$x \in \mathbb{R}^2$

$x_1 = (1,0)^T$, $y_1 = +1$

$x_2 = (-1,0)^T$, $y_2 = -1$

$x_3 = (1,d)^T$, $y_3 = +1$

(a) $d > 0$



(b) Initialize Weight Vector $\underline{w}$ To Be

$$\underline{w} = \begin{bmatrix} w_0 = 0 \\ w_1 = 0 \\ w_2 = 0 \end{bmatrix} \in \mathbb{R}^3$$

Perception Algorithm, First Round (1, 2, 3)

Point $\underline{x_1} = (1, 1, 0)^T \in \mathbb{R}^3$

$$\hat{y}_1 = \text{sign}(\underline{w^T x_1}) = \text{sign}(w_0 x_{10} + w_1 x_{11} + w_2 x_{12})$$
$$= \text{sign}((0)(1) + (0)(1) + (0)(0))$$
$$= \text{sign}(0) \neq y_1$$

$$\underline{w} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + (1) \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

Point $\underline{x_2} = (1, -1, 0)^T \in \mathbb{R}^3$

$\hat{y}_2 = \text{sign}(\underline{w^T x_2}) = \text{sign}(w_0 x_{20} + w_1 x_{21} + w_2 x_{22})$

$= \text{sign}((1)(1) + (1)(-1) + (0)(0))$

$= \text{sign}(0) \neq y_2$

$$\underline{w} = \begin{vmatrix} 1 \\ 1 \\ 0 \end{vmatrix} + (-1) \begin{vmatrix} 1 \\ -1 \\ 0 \end{vmatrix} = \begin{vmatrix} 0 \\ 2 \\ 0 \end{vmatrix}$$

Point $\underline{x_3} = (1, 1, d)^T \in \mathbb{R}^3$

$\hat{y}_3 = \text{sign}(\underline{w^T x_3}) = \text{sign}(w_0 x_{30} + w_1 x_{31} + w_2 x_{32})$

$= \text{sign}((0)(1) + (2)(1) + (0)(d))$

$= \text{sign}(2) = 1 = y_2$

No change to weight vector $\underline{w}$

Perceptron Algorithm, Second Round (1, 2, 3)

Point $\underline{x_1} = (1, 1, 0)^T \in \mathbb{R}^3$

$\hat{y}_1 = \text{sign}(\underline{w^T x_1}) = \text{sign}(w_0 x_{10} + w_1 x_{11} + w_2 x_{12})$

$= \text{sign}((0)(1) + (2)(1) + (0)(0))$

$= \text{sign}(2) = 1 = y_1$

No change to weight vector $\underline{w}$

Point $\underline{x_2} = (1, -1, 0)^T \in \mathbb{R}^3$

$\hat{y}_2 = \text{sign}(\underline{w^T x_2}) = \text{sign}(w_0 x_{20} + w_1 x_{21} + w_2 x_{22})$

$= \text{sign}((0)(1) + (2)(-1) + (0)(0))$

$= \text{sign}(-2) = -1 = y_2$

No change to weight vector $\underline{w}$

Point $\underline{x_3} = (1, 1, d)^T \in \mathbb{R}^3$

$\hat{y}_3 = \text{sign}(\underline{w^T x_3}) = \text{sign}(w_0 x_{30} + w_1 x_{31} + w_2 x_{32})$

$= \text{sign}((0)(1) + (2)(1) + (0)(d))$

$= \text{sign}(2) = 1 = y_2$

No change to weight vector $\underline{w}$

$\rho = \min_{1 \leq n \leq N} y_n(w^{*T} x_n) = 2 \qquad \underline{w^*} = \begin{vmatrix} 0 \\ 2 \\ 0 \end{vmatrix}$

$2x_1 + 0x_2 = 0$

Part C, Initialize Weight Vector $\underline{w}$ To Be

$\underline{w} = \begin{vmatrix} w_0 = 0 \\ w_1 = 0 \\ w_2 = 0 \end{vmatrix} \in \mathbb{R}^3$

Perceptron Algorithm, First Round (3, 2, 1)

Point $\underline{x_3} = (1, 1, d)^T \in \mathbb{R}^3$

$$\hat{y}_3 = \text{sign}(\underline{w^T x_3}) = \text{sign}(w_0 x_{30} + w_1 x_{31} + w_2 x_{32})$$
$$= \text{sign}((0)(1) + (0)(1) + (0)(d))$$
$$= \text{sign}(0) \neq y_1$$

$$\underline{w} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + (1) \begin{bmatrix} 1 \\ 1 \\ d \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ d \end{bmatrix}$$

Point $\underline{x_2} = (1, -1, 0)^T \in \mathbb{R}^3$

$$\hat{y}_2 = \text{sign}(\underline{w^T x_2}) = \text{sign}(w_0 x_{20} + w_1 x_{21} + w_2 x_{22})$$
$$= \text{sign}((1)(1) + (1)(-1) + (d)(0))$$
$$= \text{sign}(0) \neq y_2$$

$$\underline{w} = \begin{bmatrix} 1 \\ 1 \\ d \end{bmatrix} + (-1) \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ d \end{bmatrix}$$

Point $\underline{x_1} = (1, 1, 0)^T \in \mathbb{R}^3$

$$\hat{y}_1 = \text{sign}(\underline{w^T x_1}) = \text{sign}(w_0 x_{10} + w_1 x_{11} + w_2 x_{12})$$
$$= \text{sign}((0)(1) + (2)(1) + (d)(0))$$
$$= \text{sign}(2) = 1 = y_1$$

No change to weight vector $\underline{w}$

Point $\underline{x_3} = (1, 1, d)^T \in \mathbb{R}^3$

$\hat{y}_3 = \text{sign}(\underline{w^T x_3}) = \text{sign}(w_0 x_{30} + w_1 x_{31} + w_2 x_{32})$

$= \text{sign}((0)(1) + (2)(1) + (d)(d))$

$= \text{sign}(2 + d^2) = 1 = y_3$

No change to weight vector $\underline{w}$

Point $\underline{x_2} = (1, -1, 0)^T \in \mathbb{R}^3$

$\hat{y}_2 = \text{sign}(\underline{w^T x_2}) = \text{sign}(w_0 x_{20} + w_1 x_{21} + w_2 x_{22})$

$= \text{sign}((0)(1) + (2)(-1) + (d)(0))$

$= \text{sign}(-2) = -1 = y_2$

No change to weight vector $\underline{w}$

Point $\underline{x_1} = (1, 1, 0)^T \in \mathbb{R}^3$

$\hat{y}_1 = \text{sign}(\underline{w^T x_1}) = \text{sign}(w_0 x_{10} + w_1 x_{11} + w_2 x_{12})$

$= \text{sign}((0)(1) + (2)(1) + (d)(0))$

$= \text{sign}(2) = 1 = y_1$

No change to weight vector $\underline{w}$

$$\underline{w^*} = \begin{vmatrix} 0 \\ 2 \\ d \end{vmatrix}$$

$$\rho = \min_{1 \le n \le N} y_n(w^{*T} x_n) = 2$$

Independent of d

$2x_1 + dx_2 = 0$

$dx_2 = -2x_1$

$x_2 = -\frac{2}{d} x_1$

(d) Choose b, since it is less suceptable to noise in dataset