(1004385) Dylan Tan

1a) No such classifier exists. This is because the distances of points (1,1), (1,-1) and (-1,1) from the origin, are all the same. This means that if a circle is used as a classifier, either ALL points (1,1), (1,-1) and (-1,1) lie inside or ALL points lie outside, the circle. Hence, it would be impossible to separate (1,1), (1,-1) and (-1,1) into 2 different classes, no matter the radius r.

1b) Given points (1,1), (2,2), (-1,1) and (1,-1)

Assume that linear classifier $h(x;\theta)$ exists.

$h(x;\theta) = \text{Sign}(\theta_1 x_1 + \theta_2 x_2) = \text{sign}(\theta \cdot x) = \begin{cases} +1, & \theta \cdot x \geq 0 \\ -1, & \theta \cdot x < 0 \end{cases}$

$x^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $y^{(1)} = 1$

$x^{(2)} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, $y^{(2)} = 1$

$x^{(3)} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $y^{(3)} = -1$

$x^{(4)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $y^{(4)} = -1$

$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$

Initialize $\theta^{(1)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$y^{(1)}(\theta^{(1)} \cdot x^{(1)}) = 0 \leq 0$

update $\theta^{(1)}$:

$\theta^{(2)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + (1)\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$y^{(2)}(\theta^{(2)} \cdot x^{(2)}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}^T \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 2+2 = 4 > 0$ (no update)

$y^{(3)}(\theta^{(2)} \cdot x^{(3)}) = (-1)\begin{bmatrix} 1 \\ 1 \end{bmatrix}^T \begin{bmatrix} -1 \\ 1 \end{bmatrix} = (-1)(0) = 0 \leq 0$

update $\theta^{(2)}$:

$\theta^{(3)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + (-1)\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

$y^{(4)}(\theta^{(3)} \cdot x^{(4)}) = (-1)\begin{bmatrix} 2 \\ 0 \end{bmatrix}^T \begin{bmatrix} 1 \\ -1 \end{bmatrix} = (-1)(2) = -2 \leq 0$

update $\theta^{(3)}$:

$\theta^{(4)} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} + (-1)\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Since $\theta^{(4)} = \theta^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, repeating the process will only produce the same results, and a classifier thus does not exist.

Trg err =