ML HW4

Q1.

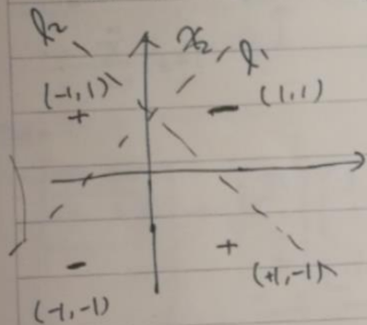
1. a.

$x_1 = 1$ $x_2 = 1$ $y = -1$ "-"

$x_1 = 1$ $x_2 = -1$ $y = 1$ "+"

$x_1 = -1$ $x_2 = -1$ $y = -1$ "-"

$x_1 = -1$ $x_2 = 1$ $y = 1$ "+"



⇒ we cannot use one perceptron to distinguish y. Thus, it cannot be used.

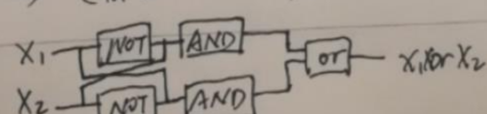
b.

x_1	x_2	x_1 (AND) (NOT(x_2))	(NOT(x_1)) (AND) x_2	x_1 or x_2
1	1	-1	-1	1
1	-1	1	-1	1
-1	1	-1	1	1
-1	-1	-1	-1	-1

c.

x_1	x_2	x_1 xor x_2
1	1	-1
1	-1	1
-1	1	1
-1	-1	-1

$$\Rightarrow (x_1 \text{ AND NOT}(x_2)) \text{ or } (\text{NOT}(x_1) \text{ AND } x_2) = \begin{cases} 1 & x_1 \neq x_2 \\ -1 & x_1 = x_2 \end{cases}$$



Q3.

$= \phi^T\underbrace{(\phi\phi^T)y}_{\alpha}.$

3 a. Because we use linear predictor w

So the ~~E~~ error $e = \Sigma(y - \phi(x)\cdot w) = \sum_{i=1}^{n}(y_i - \phi(x_i)w_i)$

b. error $= \alpha\|y - \phi w\|_2^2$ in order to make calculation easy,

let $\alpha = \frac{1}{2}$

$\therefore L(w) = \frac{1}{2}\|y - \phi w\|_2^2$

$\frac{\partial L(w)}{\partial w} = \phi^T(y - \phi w) = \phi^T y - \phi^T\phi w$

Set $\frac{\partial L(w)}{\partial w} = 0 \Rightarrow \phi^T y - \phi^T\phi w = 0$ $\quad w^* = (\phi^T\phi)^{-1}\phi^T y$

$\phi^T\phi w = \phi^T y$

$w = (\phi^T\phi)^{-1}\phi^T y$

c. $\because f(z) = \langle w, \phi(z)\rangle$ which is dot product.

$\therefore f(z) = w\cdot\phi(z) = w^*\cdot\phi(z) = (\phi^T\phi)^{-1}\phi^T y\,\phi(z)$

d. $\because (A^{-1} + B^T C^{-1} B)^{-1} B^T C^{-1} = AB^T(BAB^T + C)^{-1}$

Set $B = \phi$ $\quad C = I_n$

$\therefore$ for b $\Rightarrow$ $w^* = (\phi^T\phi)^{-1}\phi^T y = \phi^T(\phi\phi^T)^{-1}y = \frac{\Sigma\phi(x_i)y_i}{\Sigma\phi^2(x_i)} = \frac{\sum_{i=1}^{n}k(x_i, y_i)}{\sum_{i=1}^{n}k(x_i, x_i)}$ $\quad k = \phi\phi^T$

for c $\Rightarrow$ $f(z) = \phi^T(\phi\phi^T)\cancel{y}\phi(z)y = \frac{\Sigma\phi(x_i)\phi(z)y_i}{\Sigma\phi^2(x_i)} = \frac{\sum_{i=1}^{n}k(x_i, z)y_i}{\sum_{i=1}^{n}k(x_i, x_i)}$ $\quad k = \phi\phi^T$