Kwong Yuet Michael Fadillah Wong

1a

$$\prod_{i=1}^{n} p(y\ |X = \ x_i; w, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(y_i - wx_i)^2}{2\sigma^2}}$$

1b

$$f(\mathrm{w,y}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(y_i - wx_i)^2}{2\sigma^2} * \left(\frac{1}{\sqrt{2\pi p^2}}\right) e^{-\frac{w^2}{2p^2}}}$$

2

$$\text{Let Y be } (y_i - \hat{y})^2, X\ be\ (1, x_i, x_i^3), W = (w_0, w_1, w_2)$$

$$\frac{\Delta \mathrm{Y}}{\Delta W} = \left(\frac{\Delta Y}{\Delta y_i}\right) * \left(\frac{\Delta y_i}{\Delta \mathrm{W}}\right) = -2(y_i - \ \hat{y}_i)X = -2(y_i - w_0 - w_0 x_i - 2w2x_i)X$$

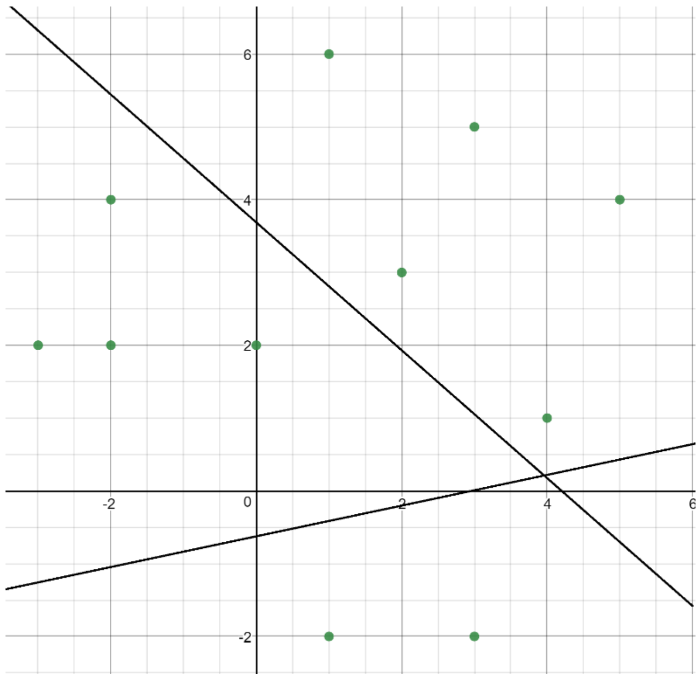
$$w_0 = \ w_0 + 2(y_i - w_0 - w_1 x_i - w^2 x_i^3) = 2y_i - w_0 - 2w_1 x_i - 2w2x_i^3$$

$$w_1 = \ w_1 + 2(y_i - w_0 - w_1 x_i - w^2 x_i^3)x_i = 2y_i x_i - 2w_0 x_i + (1 - 2x_i^2)w_1 - 2w2x_i^4$$

$$w_2 = \ w_2 + 2(y_i - w_0 - w_1 x_i - w^2 x_i^3)x_i^3 = 2y_i x_i^3 - 2w_0 x_i - 2w_1 x_i^2 + (1 - 2x_i^6)w_2$$

$$\hat{y}_i = 2y_i - w_0 - 2w_1 x_i - 2w_2 x_i^3 + \left(2y_i x_i - 2w_0 x_i + (1 - 2x_i^2)w_1 - 2w_2 x_i^4\right)x_i \\ + \left(2y_i x_i^3 - 2w_0 x_i - 2w_1 x_i^2 + (1 - 2x_i^6)w_2\right)x_i^3$$

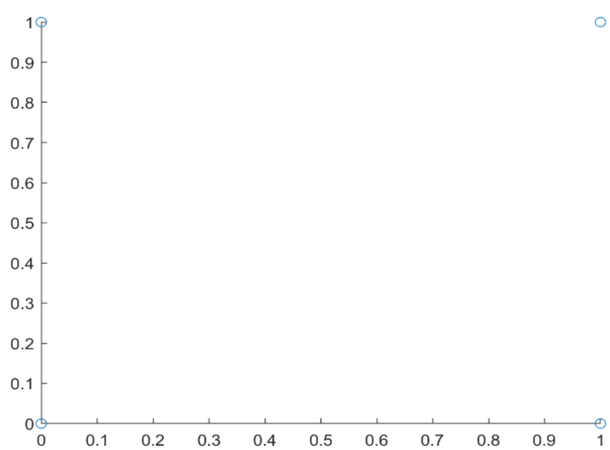
3a Since there are three inputs and no hidden layer in this network, classification problem like "xor" cannot be solved through this network. By this plot, we can see that it is possible to classify them by a straight line. Therefore, the problem can be solved.



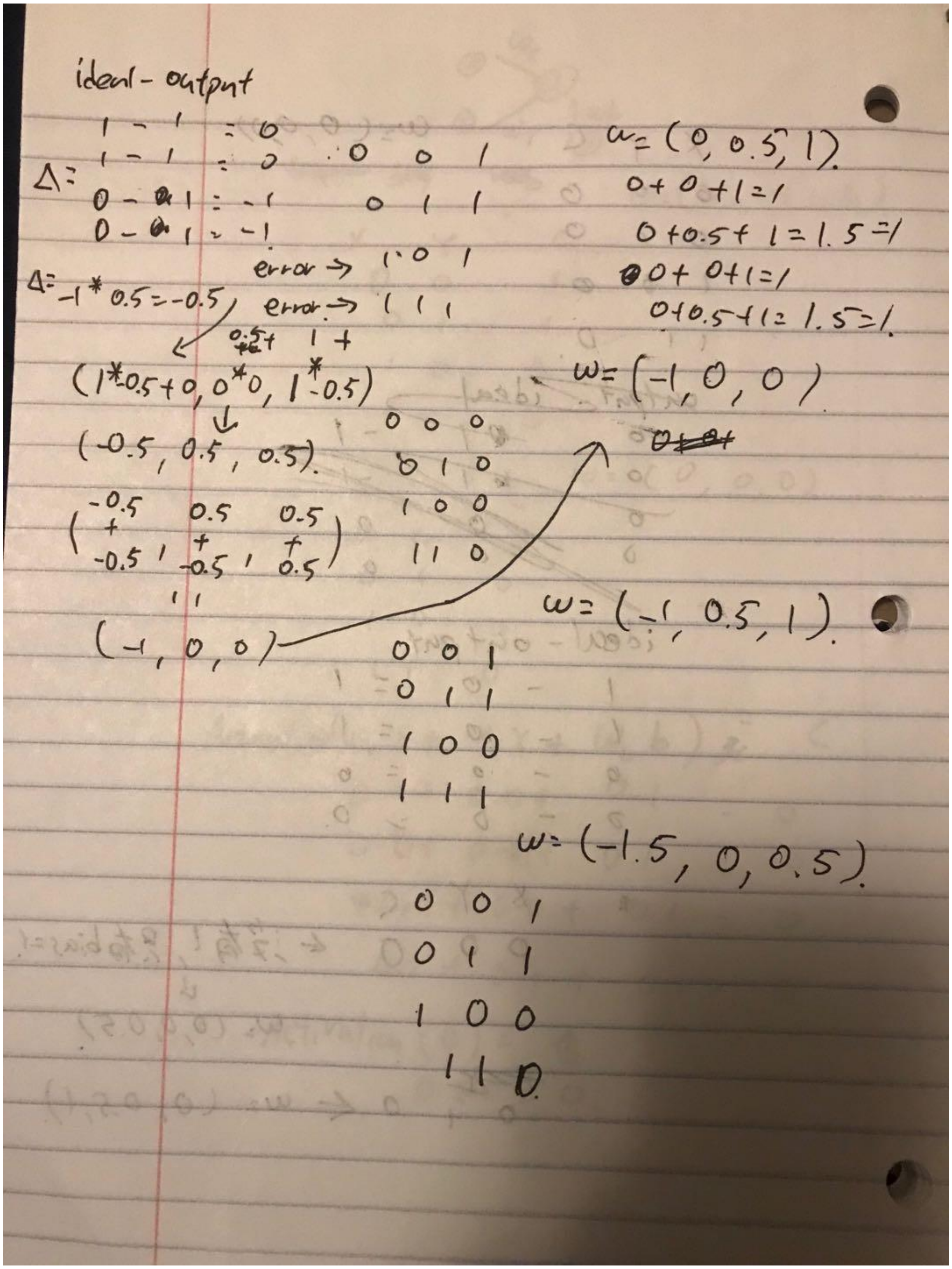
3b

If we add (-1,6) to C1, we cannot learn the sample through this network, because we are not able to separate them with one straight line. In other word, a hidden layer is needed in order to solve the problem.

4a



Here are some partial steps that I did for question 4b.



ideal - output

$1 - 1 = 0$

$1 - 1 = 0$ 0 0 1

Δ:

$0 - 1 = -1$ 0 1 1

$0 - 1 = -1$

error → 1 0 1

$\Delta = -1 * 0.5 = -0.5$ error → 1 1 1

0.5+ 1+

$(1*0.5+0, 0*0, 1*-0.5)$

$(-0.5, 0.5, 0.5)$

$(-0.5 + -0.5, 0.5 + -0.5, 0.5 + 0.5)$

$(-1, 0, 0)$

w = (0, 0.5, 1)

0+0+1=1

0+0.5+1=1.5=1

0+0+1=1

0+0.5+1=1.5=1

w = (-1, 0, 0)

~~0+0+~~

0 0 0

0 1 0

1 0 0

1 1 0

w = (-1, 0.5, 1)

0 0 1

0 1 1

1 0 0

1 1 1

w = (-1.5, 0, 0.5)

0 0 1

0 1 1

1 0 0

1 1 0

4b

W = (Wx_1, Wx_2, Wb)

W = (0,0,0) -> (0, 0, 0.5) -> (0,0.5,1) -> (-0.5, 0.5, 0.5) -> (-1,0,0) -> (-1,0.5,0.5) -> (-1,0.5,1) -> (-1.5,0,0.5)

Final weight = (-1.5,0,0.5)

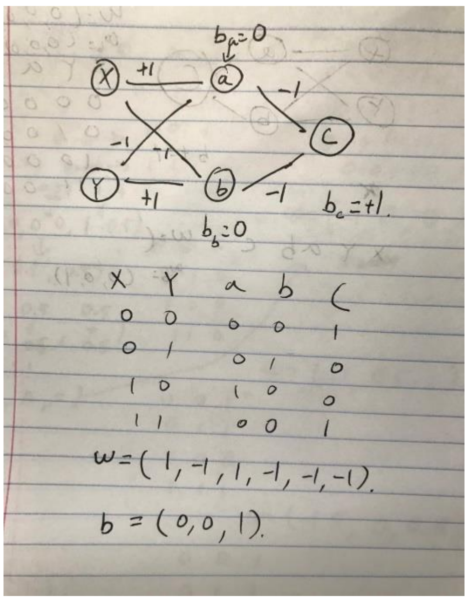
4c

Predict $y = 1 \; if \; -1.5x_1 + 0.5 > 0$

Predict $y = 0 \; if \; -1.5x_1 + 0.5 \leq 0$

4db

Since it is not possible to divide them with a straight line, we need to add a layer in order to compute the result.



4dc

Predict $y = 1 \; if \;\; x_{11} - x_{12} + x_{22} - x_{21} - x_{31} - x_{32} + 1 > 0$

Predict $y = 0 \; if \;\; x_{11} - x_{12} + x_{22} - x_{21} - x_{31} - x_{32} + 1 \leq 0$

5a

$$\frac{\mathrm{dE}}{dw_j(n)} = \frac{1}{2}\sum_{i=1}^{n} 2(d(i) - (e^{-\frac{1}{2\sigma^2(n)}\left||x_i - u_j(n)|\right|^2})$$

$$\frac{\mathrm{dE}}{du_j(n)} = = \frac{1}{2}\sum_{i=1}^{n} 2(d(i) - (w_j(n)e^{-\frac{1}{2\sigma^2(n)}2||x_i - 1||})$$

$$\frac{\mathrm{dE}}{d\sigma(n)} = \frac{1}{2}\sum_{i=1}^{n} 2(d(i) - (w_j(n)e^{-\frac{1}{4}\left||x_i - u_j(n)|\right|^2})$$

5b

$$\frac{1}{2}\sum_{i=1}^{n} 2(d(i) - (e^{-\frac{1}{2\sigma^2(n)}\left||x_i - u_j(n)|\right|^2}) * learningrate_{wj(n)} * \frac{1}{n}$$

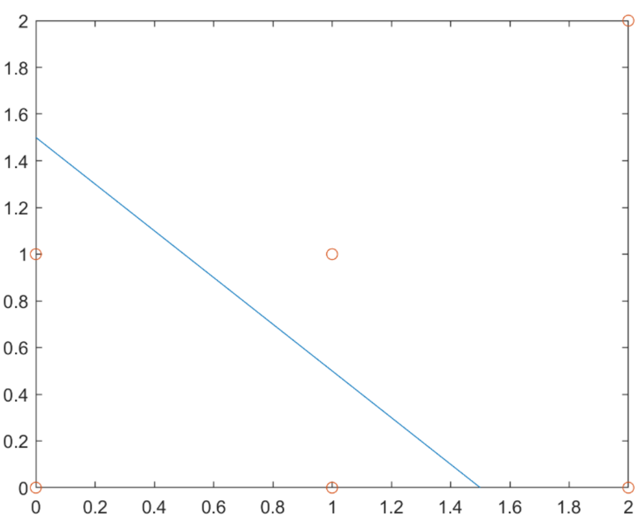
$$\frac{1}{2}\sum_{i=1}^{n} 2(d(i) - \left(w_j(n)e^{-\frac{1}{2\sigma^2(n)}2\left||x_i - 1|\right|}\right) * learningrate_{uj(n)} * \frac{1}{n}$$

$$\frac{1}{2}\sum_{i=1}^{n} 2(d(i) - (w_j(n)e^{-\frac{1}{4}\left||x_i - u_j(n)|\right|^2}) * learningrate_{\sigma(n)} * \frac{1}{n}$$

6a

Weight vector: $[2,2]^T, b = -3$

Optimal margin $= \frac{\sqrt{2}}{2}$



6b

Support vectors: (0,1), (1,0), (1,1), (2,0)