HW 7

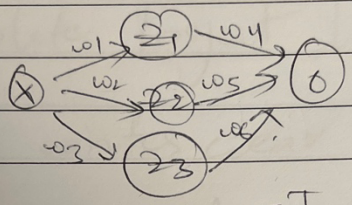
1.1

Date: / /

HW7

1. Exercises

1.1 (a)



$w = [1, 1, -1, 0.5, 1, 2]^T$

$x = 4$

$z_1 = w_1 x = 1 \times 4 = 4$

$z_2 = w_2 x = 4$

$z_3 = w_3 x = -4.$

For hidden layer z_1, z_2, z_3

$z = z_1 w_4 + z_2 w_5 + z_6 w_6$

$= 2 + 4 + 0$

$= 6.$

MATRIKAS

Date: / /

For o/p node, Sigmoid activation fn.

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

$$= \frac{1}{1+e^{-2}} = \frac{1}{1+e^{-6}}$$

b) $= 0.995$

$$\text{loss} = (y-\hat{y})^2$$

$$= (y-a)^2$$

$$= (0-0.995)^2$$

$$\text{loss} = 0.9950$$

c) $\frac{dL}{da} = -2(y-a)$

$$= 2(0.9975)$$

$$= 1.9950$$

$$\frac{da}{dz} = d(\sigma(z))$$

$$= a(1-a)$$

$$= 0.0024$$

MATRIKAS

Date: / /

$$\frac{dl}{dz} = \frac{dl}{da} \times \frac{da}{dz}$$

$$= 1.9950 \times 0.0025$$

$$= 0.0050$$

$$\frac{dl}{d\omega_4} = \frac{dl}{dz} \times \frac{dz}{d\omega_4}$$

$$= 0.0050 \times 4$$

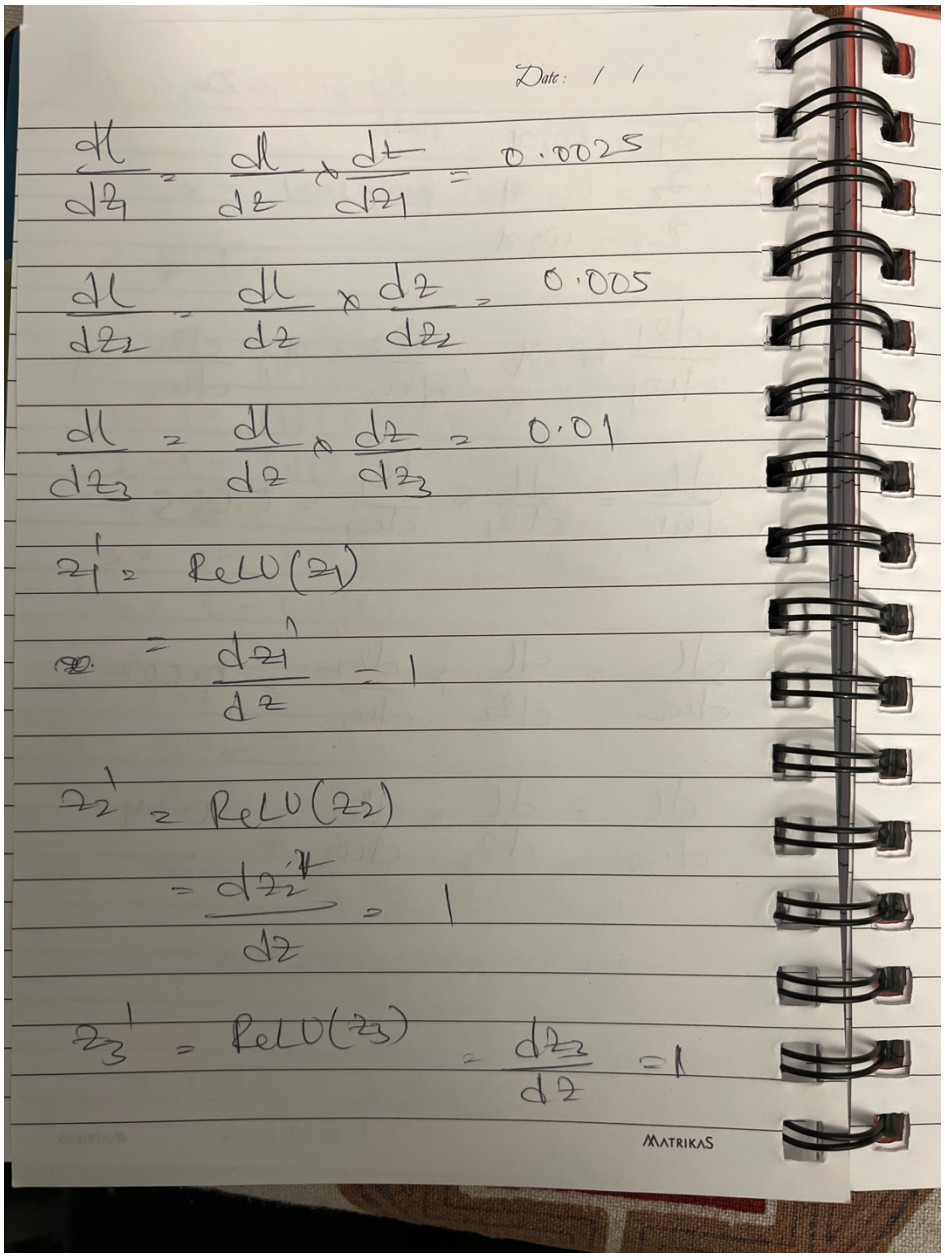
$$= 0.02$$

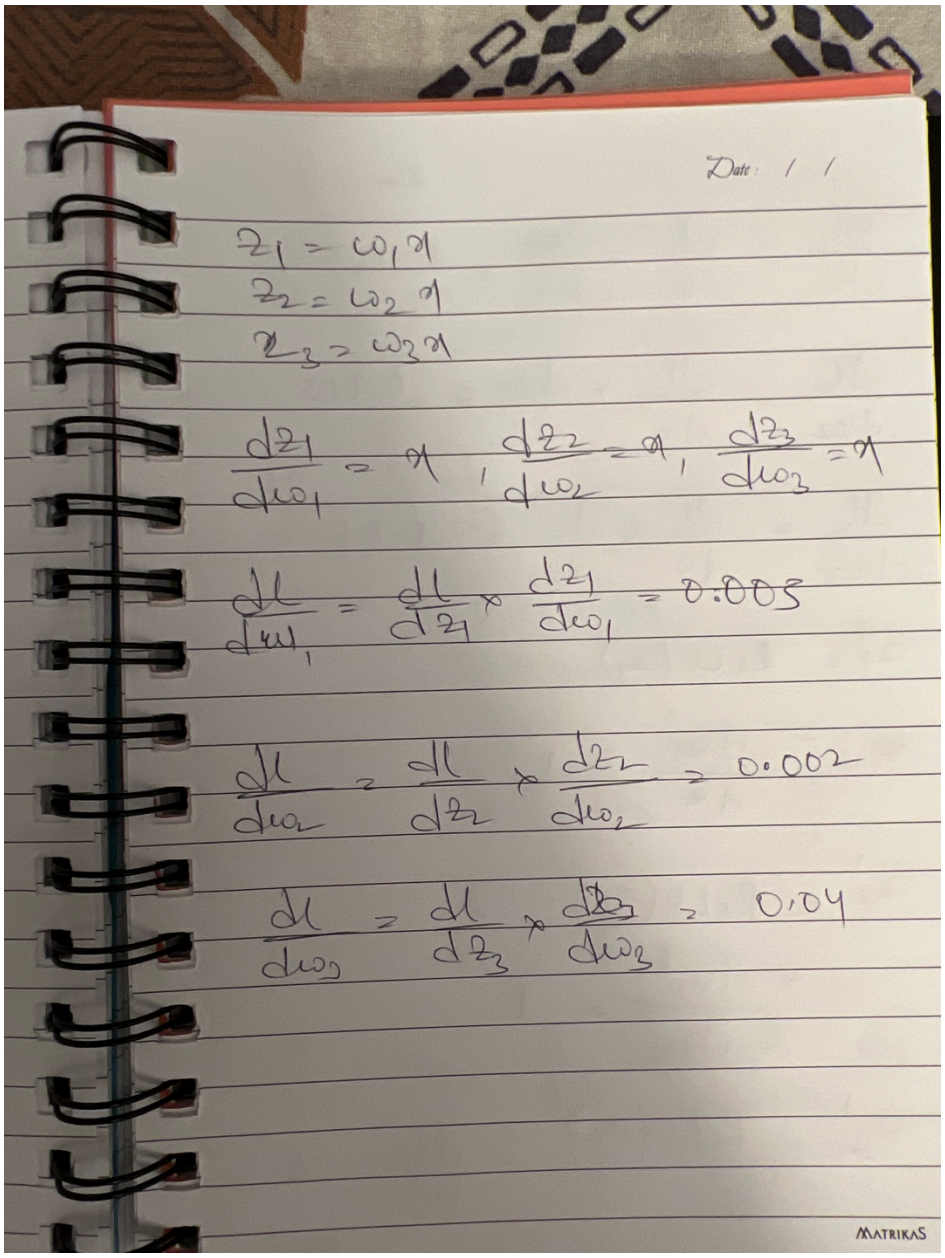
$$\frac{dl}{d\omega_5} = \frac{dl}{dz} \times \frac{dz}{d\omega_5} = 0.02$$

$$\frac{dl}{d\omega_6} = \frac{dl}{dz} \times \frac{dz}{d\omega_6} = 0.$$

$$\frac{dl}{dz_1} = \omega_4, \quad \frac{dl}{dz_2} = \omega_5, \quad \frac{dl}{dz_3} = \omega_6.$$

MATRIKAS


Date: / /
$\frac{dL}{dz_4} = \frac{dL}{dz} \times \frac{dz}{dz_1} = 0.0025$
$\frac{dL}{dz_2} = \frac{dL}{dz} \times \frac{dz}{dz_2} = 0.005$
$\frac{dL}{dz_3} = \frac{dL}{dz} \times \frac{dz}{dz_3} = 0.01$
$z_1' = ReLU(z_1)$
$= \frac{dz_1'}{dz} = 1$
$z_2' = ReLU(z_2)$
$= \frac{dz_2'}{dz} = 1$
$z_3' = ReLU(z_3) = \frac{dz_3}{dz} = 1$
MATRIKAS



Date: / /

$z_1 = \omega_1 x$

$z_2 = \omega_2 x$

$z_3 = \omega_3 x$

$\frac{dz_1}{d\omega_1} = x$, $\frac{dz_2}{d\omega_2} = x$, $\frac{dz_3}{d\omega_3} = x$

$\frac{dl}{d\omega_1} = \frac{dl}{dz_1} \times \frac{dz_1}{d\omega_1} = 0.005$

$\frac{dl}{d\omega_2} = \frac{dl}{dz_2} \times \frac{dz_2}{d\omega_2} = 0.002$

$\frac{dl}{d\omega_3} = \frac{dl}{dz_3} \times \frac{dz_3}{d\omega_3} = 0.04$

MATRIKAS

d) $w_1 = w_1 - \alpha \frac{dl}{dw_1}$

α = learning rate

$\alpha = 1.0$

$w_1 = 1 - (1.0)(0.01) = 0.99$

$w_2 = 1 - (1)(0.02) = 0.98$

$w_3 = -1 - (1)(0.04) = -0.96$

$w_4 = 0.5 - (0.2) = 0.46$

$w_5 = 1 - 1(0.02) = 0.98$

$w_6 = 2 - 0.1(0) = 2$

$z_1 = x w_1 = 4 \times 0.99 = 3.96$

$z_2 = x w_2 = 4 \times 0.98 = 3.92$

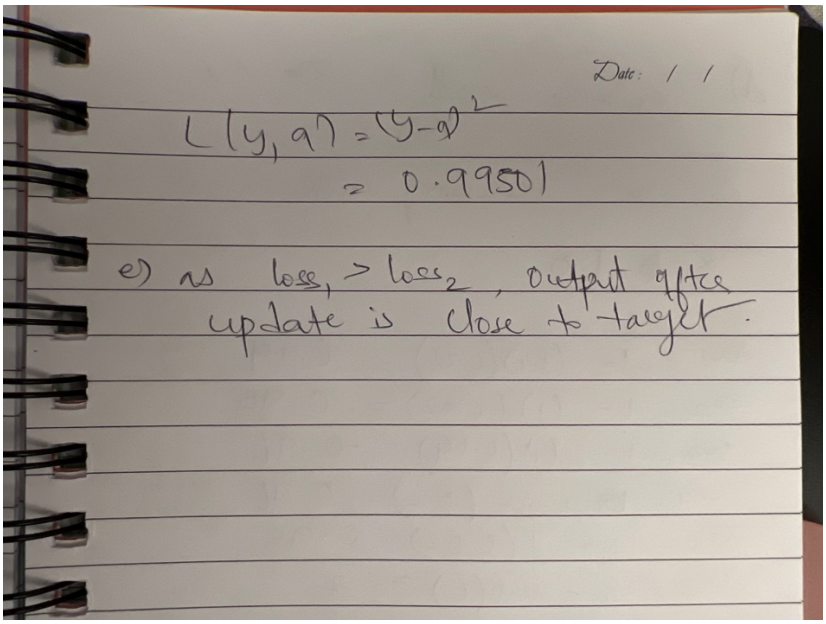
$z_3 = x w_3 = 4 \times -0.96 = -3.84$

$z_2 = z_1 w_4 + z_3 w_5 + z_3 w_6$

$= 3.96(0.46) + 3.92(0.98) + 3.84(2)$

$= 1.82 + 3.841 + 7.68$

$= 13.34$



1.2

14Q)

a. linearly separable
b. linearly separable
c. linearly separable
d. not linearly separable

15Q)

a. Lets say x,y are Boolean variables and Z is the output.

The perceptron model for AND Function is

Z = sgn(x+y-1.5)

OR Function is

Z = sgn(x+y-0.5)

b. The disadvantage of using linear functions as activation functions for multilayer neural networks is that linear function's result is always linear so having a multilayer network would be same as a linear network.