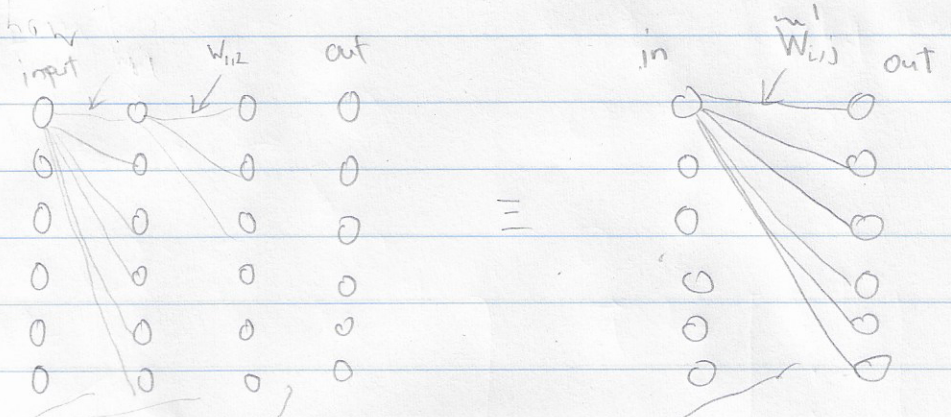
2.

Show



$$\vec{a}^{(1)} = W^{(1)}\vec{a}^{(0)} + \vec{b}^{(1)} \qquad \vec{a}^{(1)} = W^{(1)}\vec{a}^{(0)} + \vec{b}^{(1)}$$

$$\vec{a}^{(2)} = W^{(2)}\vec{a}^{(1)} + \vec{b}^{(2)}$$

$$\vec{a}^{(3)} = W^{(3)}\vec{a}^{(2)} + \vec{b}^{(3)}$$

output

$$\Rightarrow \vec{a}^{(3)} = W^{(3)}\left[W^{(2)}\vec{a}^{(1)} + \vec{b}^{(2)}\right] + \vec{b}^{(3)}$$

$$= W^{(3)}\left[W^{(2)}\left(W^{(1)}\vec{a}^{(0)} + \vec{b}^{(1)}\right) + \vec{b}^{(2)}\right] + \vec{b}^{(3)}$$

$$= W^{(3)}\left[W^{(2)}W^{(1)}\vec{a}^{(0)} + W^{(2)}\vec{b}^{(1)} + \vec{b}^{(2)}\right] + \vec{b}^{(3)}$$

$$\vec{a}^{(3)} = W^{(3)}W^{(2)}W^{(1)}\vec{a}^{(0)} + W^{(3)}W^{(2)}\vec{b}^{(1)} + W^{(3)}\vec{b}^{(2)} + \vec{b}^{(3)}$$

For network 1 $\equiv$ network 2

inputs : Same inputs $\Rightarrow \vec{a}^{(0)} = \tilde{\vec{a}}^{(0)}$

Outputs equal $\Rightarrow \vec{a}^{(3)} = \tilde{\vec{a}}^{(3)}$

$$\therefore \underbrace{W^{(3)}W^{(2)}W^{(1)}\vec{a}^{(0)}}_{\vec{V}_1} + \underbrace{W^{(3)}W^{(2)}\vec{b}^{(1)}}_{\vec{V}_2} + \underbrace{W^{(3)}\vec{b}^{(2)}}_{\vec{V}_3} + \vec{b}^{(3)} = \tilde{W}^{(1)}\tilde{\vec{a}}^{(0)} + \tilde{\vec{b}}^{(1)}$$

Since, we are given network 1's weights and biases compute $\vec{V}_1$, $\vec{V}_2$, and $\vec{V}_3$.

$$\Rightarrow \tilde{W}^{(1)}\vec{a}^{(0)} + \tilde{\vec{b}}^{(1)} - \vec{V}_1 - \vec{V}_2 - \vec{V}_3 = \vec{b}^{(3)}$$

let $\vec{k} = (-\vec{V}_1 - \vec{V}_2 - \vec{V}_3)$

$$\Rightarrow \tilde{W}^{(1)}\vec{a}^{(0)} + \tilde{\vec{b}}^{(1)} + \vec{k} = \vec{b}^{(3)}$$

Since, $\tilde{\vec{a}}^{(0)} = \vec{a}^{(0)}$, we have

$$\tilde{W}^{(1)}\vec{a}^{(0)} + \tilde{\vec{b}}^{(1)} = \underbrace{\vec{b}^{(3)} - \vec{k}}_{\text{we can compute this}}$$

let the $\vec{Q}$ be an orthogonal vector of $\tilde{W}^{(1)}\vec{a}^{(0)}$

$$\Rightarrow \vec{Q}(\tilde{W}^{(1)}\vec{a}^{(0)} + \tilde{\vec{b}}^{(1)}) = \vec{b} = \vec{Q}(\vec{b}^{(3)} - \vec{k})$$

Q.E.D

$\therefore$ W (next page)

7.

7-1

2D grid, m cols, n rows

$$\begin{array}{cccccc} 0 & 1 & 2 & & & m-1 \\ (0,0), & (1,0) & (2,0) & \cdots & & (m-1,0) \\ (0,1) & & & & & \end{array}$$

a) $1D \leftarrow 2D$

$\forall$ $n \times m$ grids, we always know 2 thing, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ maps to 1

and $\begin{bmatrix} 0 \\ m \end{bmatrix}$ maps to m

$\therefore$ 2D to 1D can be easily computed

let $A = 1 \times 2$ matrix $= [a_1 \quad a_2]$

$A\begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 \Rightarrow a_1 = 1$

$A\begin{bmatrix} 0 \\ 1 \end{bmatrix} = m \Rightarrow a_2 = m$

$\Rightarrow A = [1 \quad m] \quad \therefore f: 2D \to 1D = X_1 + mX_2$

recall $X_1 + mX_2 = C = R + mX_2$, $\Rightarrow$ $1 \leq R < m$

now, $1D \rightarrow 2D$, let $c \in 1D$

we can check for remainder to compute x_1

$c \% m = x_1$

$\left\lfloor \frac{c}{m} \right\rfloor = X_2$