the sigmoid nonlinearity.

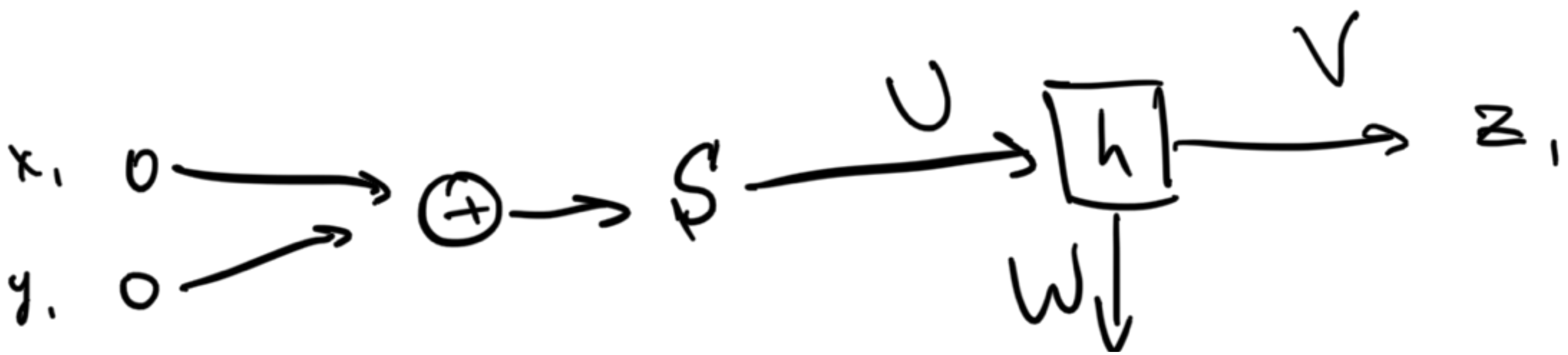
3)

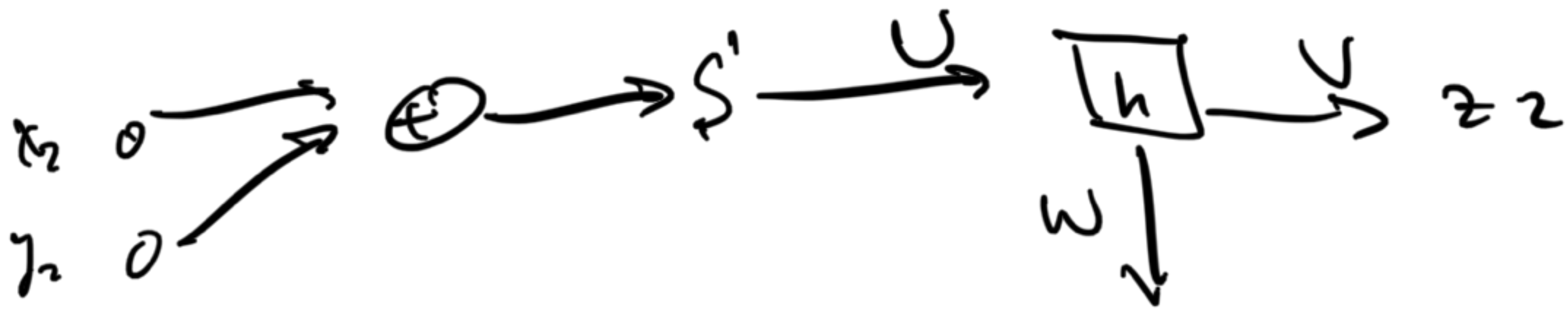
101 + 001 = 110

In 1: [1, 0, 1, 0]

In 2: [0, 0, 1, 0]

out = [0, 1, 1, 0]





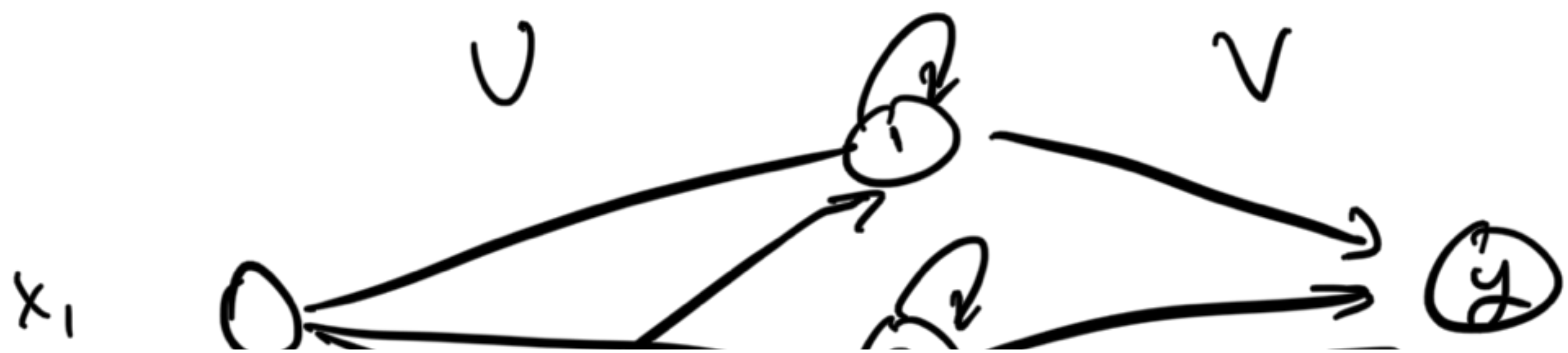
OUTPUT weight V activates when:

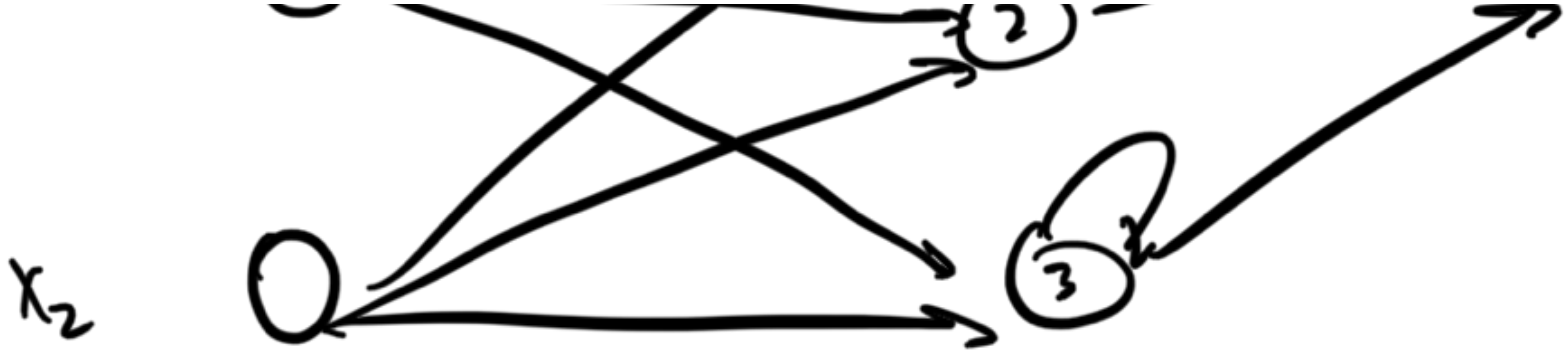
NT: each hidden unit is a hyperplane

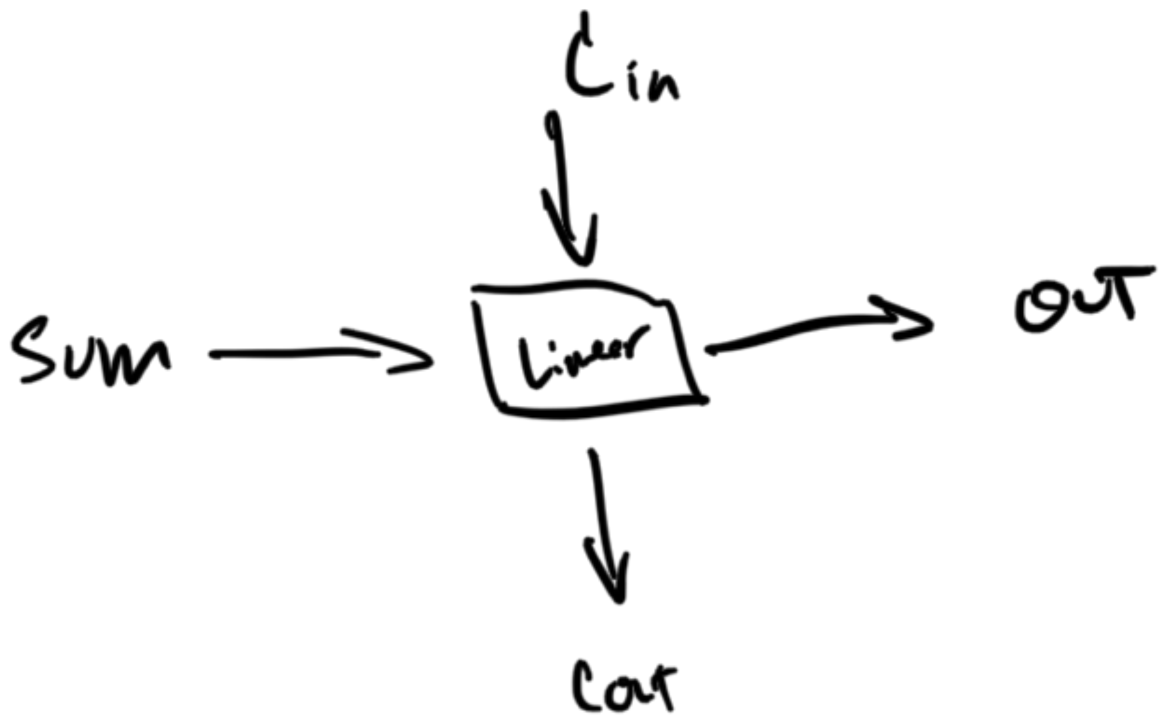
x	y	s
0	0	0
0	1	1
1	0	1
1	1	10

W activates when carry 1

U activates when:







We want Linear to separate 3 cases

Sum = 0 OUT ≥ 0 COUT ≤ 0

Cout ≤ 0

sum = 1

$out = \{$ ... out

1 1

= 2

$W_{out} = 1$

sum > 0

OUT

sum > 1

$$OUT: \begin{matrix} 0 \\ 1 \\ 1 \end{matrix} = \underline{\quad} \begin{matrix} 0 \\ 1 \\ 2 \end{matrix} + (\underline{-1})$$

$$= \underline{\quad} \begin{matrix} -1 \\ 0 \\ 1 \end{matrix}$$

$$OUT = sign\left(\underset{w}{1} \cdot \underset{sum}{x} \underset{bias}{-1}\right)$$

$x = 0 \rightarrow -1 \rightarrow sign = 0$

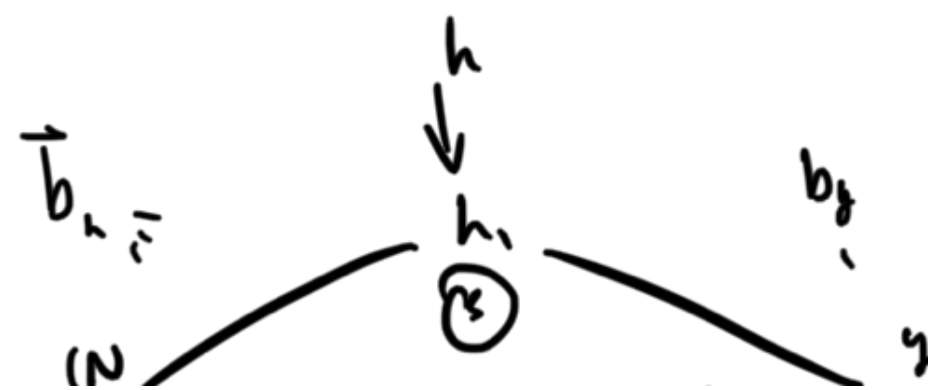
$x = 1 \rightarrow 0 \rightarrow sign = 1$

$x = 2 \rightarrow 1 \rightarrow \text{sign} = 1$

But C_{in}!

S	C_{out}	OUT
0	0	0
1	0	1
2	1	0
3	1	1

This is why we need to use hidden state to determine output



$$\underset{y}{\text{OUT}} = \overset{1\times 3}{V}\overset{3\times 1}{\vec{h}} + b_y$$

$$\overset{3\times 3}{W}\overset{3\times 1}{\vec{h}} + \overset{3\times 1}{\vec{b}_h} + \overset{3\times 2}{U}\overset{2\times 1}{\vec{x}}$$

$h_{t+1} = W x_{in} + b_h$

We want h_{t-1} to encode 4 states

$\forall_h$, out =

$$
\begin{array}{cccc}
 & \geq 3 & \geq 2 & \geq 1 \\
0 \leftarrow & [\,0, & 0, & 0\,] = 0 \\
1 \leftarrow & [\,0, & 0, & 1\,] = 1 \\
0 \leftarrow & [\,0, & 1, & 1\,] = 2 \\
1 \leftarrow & [\,1, & 1, & 1\,] = 3
\end{array}
$$

$$y_{out} = \text{sign}(V\vec{h} + b) = \text{sign}\left(\begin{bmatrix}2\\-2\\2\end{bmatrix}\cdot\vec{h} - 1\right)$$

$$[2\ -2\ 2]\begin{bmatrix}0\\0\\0\end{bmatrix} - 1 = -1 \Rightarrow 0$$

$$[2\ -2\ 2]\begin{bmatrix}0\\0\\1\end{bmatrix} - 1 = 1 \Rightarrow 1$$

$$[2\ -2\ 2]\begin{bmatrix}0\\1\\1\end{bmatrix} - 1 = -1 \Rightarrow 0$$

$$[2\ -2\ 2]\begin{bmatrix}1\\1\\1\end{bmatrix} - 1 = 1 \Rightarrow 1$$

$$W = [2, -2, 2] \qquad b = -1$$

HIDDEN OUTPUT

New hidden: $h_t = W\vec{h}_{t-1} + \vec{b}_h + \vec{U}x_t$

$$
\underset{h_t}{\begin{matrix} [\quad\quad] \\ [\quad\quad] \\ [0\ 1\ 1] \\ [1\ 1\ 1] \end{matrix}}
=
\underset{\substack{W \\ 3\times 3}}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}
\underset{h_{t-1}}{\begin{matrix} [0,0,0] \\ [0,0,1] \\ [0,1,1] \\ [1,1,1] \end{matrix}}
+
\underset{\vec{b}_h}{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}}
+
\underset{\substack{U \\ 3\times 2}}{\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}}
\underset{\substack{x_t \\ 2\times 1}}{\begin{matrix} [0,0] \\ [0,1] \\ [1,0] \\ [1,1] \end{matrix}}
$$

Pick a U

$$
\begin{bmatrix} 0 \\ 0 \end{bmatrix} \to \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \checkmark
\qquad
\begin{bmatrix} 0 \\ 1 \end{bmatrix} \to \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}
\qquad
\begin{bmatrix} 1 \\ 0 \end{bmatrix} \to \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}
\qquad
\begin{bmatrix} 1 \\ 1 \end{bmatrix} \to \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \checkmark
$$

$$U = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{matrix} [0,0,0]^{h_0} \\ [0,0,1]^{h_1} \\ [0,1,1]^{h_2} \\ [1,1,1]^{h_3} \end{matrix} = \left(\overset{W}{\begin{bmatrix} 0 & & \\ & 1 & \\ & & 0 \end{bmatrix}} \overset{h}{\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}} + \overset{b}{\begin{bmatrix} \\ \\ \end{bmatrix}} \right) + \begin{matrix} [0,0,0]^T \\ [0,0,1]^T \\ [0,1,0]^T \\ [0,1,1]^T \end{matrix}$$

$$\downarrow$$

$$[0,0,0]^{h_0} = [0,0,0 \quad] + [0,0,0]$$

$[0,0,0] = [0, \dots] + [\dots]$

$[0,0,1]^{h_1} = [0,0,1] + [0,0,0]$

$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$

$[0,0,1] = [0,0,0] + [0,0,1]$

$[0,1,1] = [0,1,0] + [0,0,1]$

$(0,1,1] = [0,1,1] + [0,1,0]$

	s	h	h_{t+1}
	0	0	0
(0,0)	0	1	1
	0	2	1
	0	3	1
	0		
	1	0	1
(0,1)	1	1	1
(1,0)	1	2	2
	1	3	2

$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$

$(0,0) = 0,0,0$

$(0,1) = 1,1,1$

$(1,0) = 1,1,1$

$(1,1) = 2,2,2$

$h_{t-1} = 2 \qquad S = 2$

$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -2 \\ -2 \end{bmatrix}\right) \qquad \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

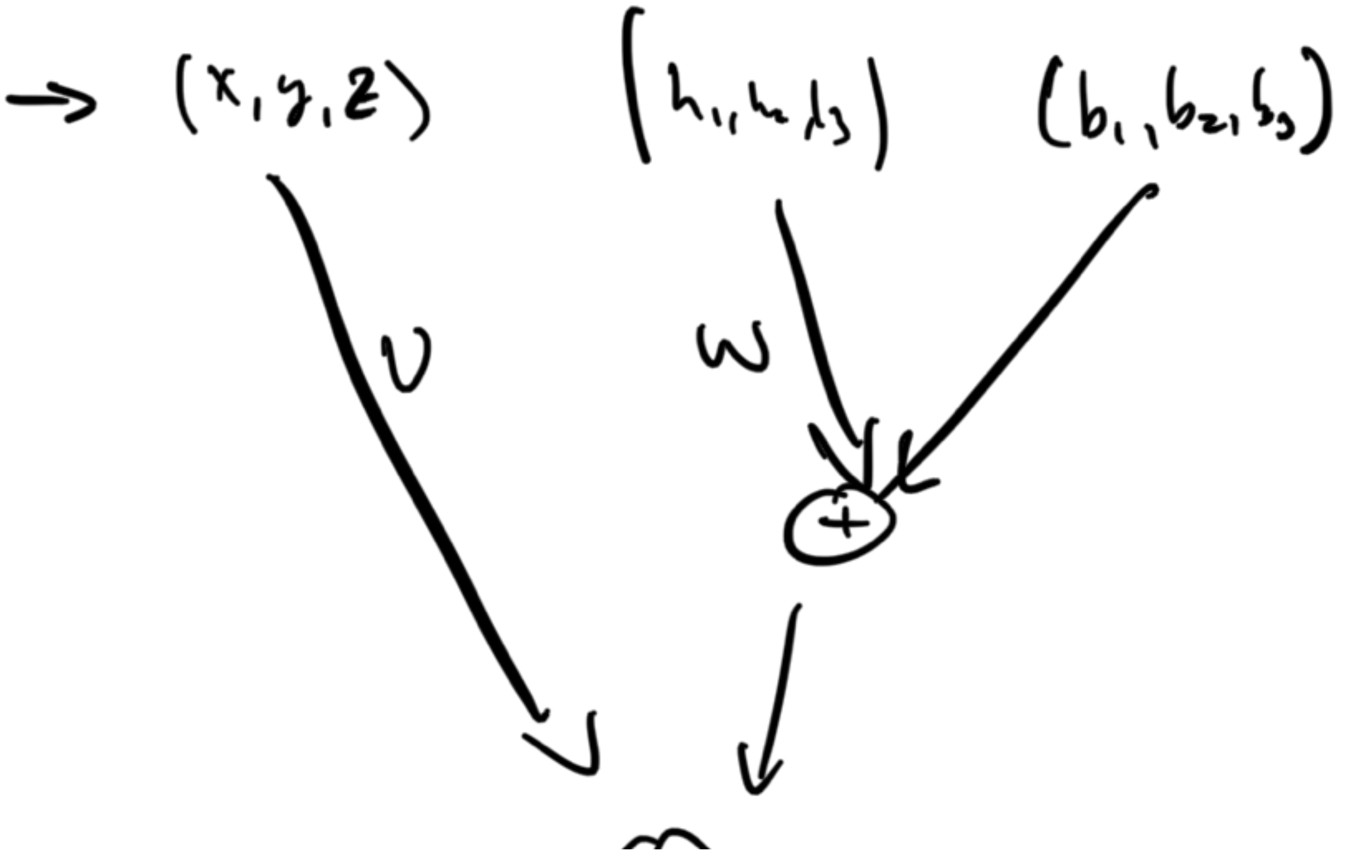
	1	0	2
(1,1)	2	1	2
	2	2	3
	2	3	3

$$\begin{bmatrix}0\\1\\1\end{bmatrix} = I \cdot \begin{bmatrix}0\\1\\1\end{bmatrix} + 0 + \begin{bmatrix}1\\1\\1\end{bmatrix}$$

x

0 , 0
0 , 1
1 , 0
1 , 1

$\vec{h}$ $\quad$ $\vec{b}_h$



$$\oplus$$

$$h$$

$$h = \mathrm{sig}\left(W h_{t-1} + \vec{b}_h\right) + Ux$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = W \cdot h + b \quad + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \quad & \quad \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$$

h should encode the sum

$$\begin{array}{lcccc} x_1 & 0 & 0 & 1 & 1 \\ x_2 & 0 & 1 & 0 & 1 \end{array} \to \overbrace{U}^{\text{new sum}} \to + \overbrace{b + Wh}^{\text{prev sum}} \to h$$

$$Ux =$$

$$[0, 0, 0]$$
$$[0, 0, 1]$$
$$[0, 0, 1]$$

$$U = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$$

$$[0, 0, 2]$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = A \cdot \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

$$U_x = \begin{array}{c} 0 \\ 1 \\ 2 \end{array} \begin{bmatrix} 1, & 0, & 0 \\ -1, & 1, & 0 \\ 0, & 0, & 0 \end{bmatrix}$$

0, 0, 0
-1, 1, 0
1, 1, 0
0, 2, 0

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}$$

0, 0
0, 1
1, 0
1, 1

$$U = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix} \qquad Ux = \underset{0}{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}}, \underset{1}{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}}, \underset{1}{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}}, \underset{2}{\begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}}$$

$$V = [2, -2, 2]$$

$$h = (W h_{t-1} + b) + Ux$$

1

$$\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \qquad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix},$$

$$\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

ⓒ

$$\begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

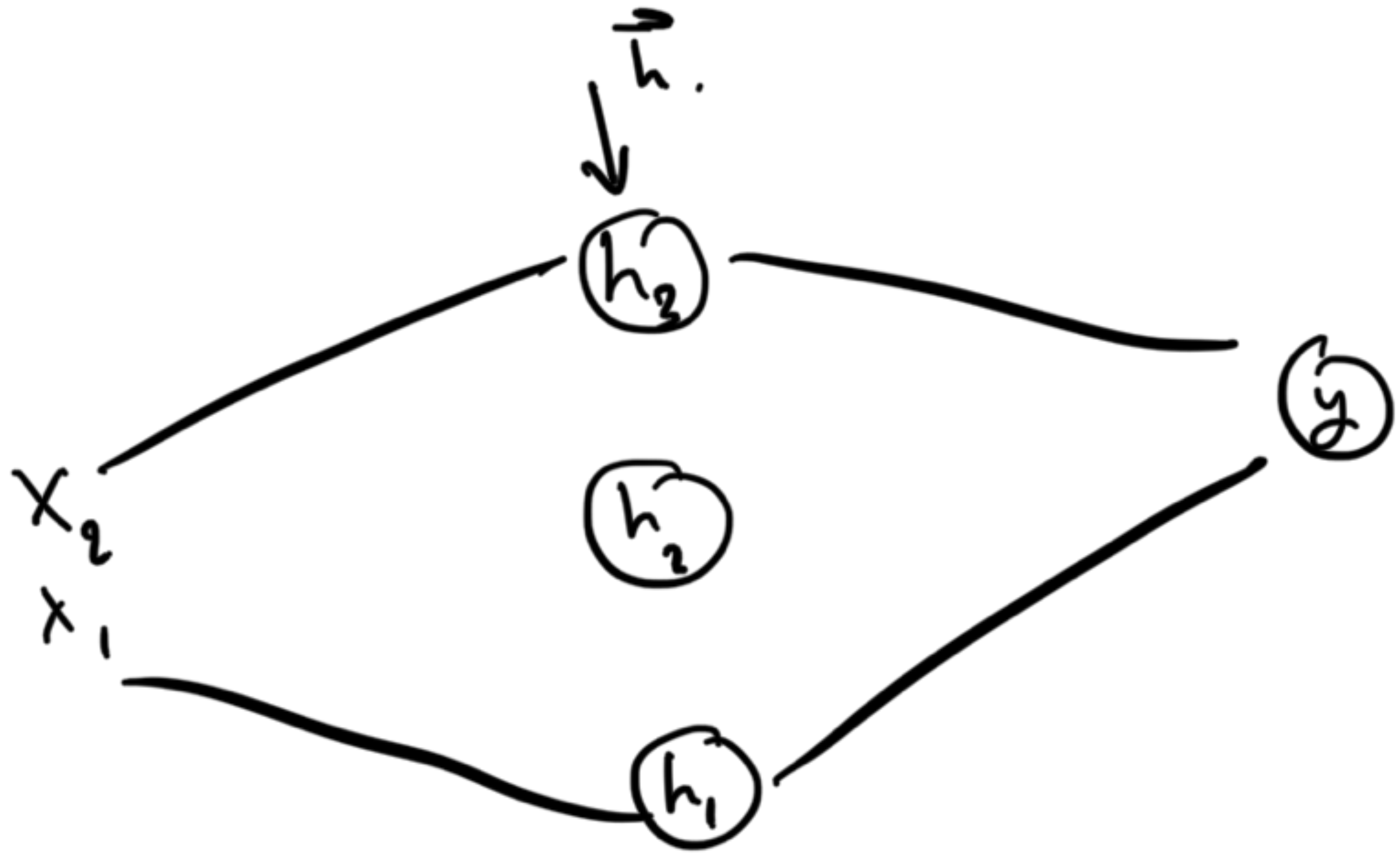
$$\begin{matrix} 1, & 0, & 1 \\ 0, & 1, & 1 \\ 0, & 0, & 1 \\ 1, & 1, & 1 \end{matrix} \rightarrow \begin{matrix} 1, & 0, & 0 \\ 0, & 1, & 0 \\ 0, & 0, & 0 \\ 0, & 0, & 1 \end{matrix}$$

$$\begin{bmatrix} & \\ & \\ & \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{aligned} 0,0,1 &\longrightarrow 0\ 0\ 0 \\ 0,1,1 &\longrightarrow 0\ 0\ 1 \\ 1,0,1 &\longrightarrow 0\ 1\ 0 \\ 1,1,1 &\longrightarrow 1\ 0\ 0 \end{aligned}$$

Each hidden unit can have its own transformation

$\geq 1 = u_1 \quad x$



x_1	x_2	s
0	0	0
0	1	1
1	0	1
1	1	2

$h_1 \longrightarrow 1 \ \forall /_{\text{except}} \ s=0$

$h_2 \longrightarrow 1$

x_1	x_2	h_3	h_2	h_1	h_3	h_2	h_1	y
0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0
0	0	0	1	1	0	0	1	1
0	0	1	1	1	0	0	1	1
0	1	0	0	0	0	0	1	1
0	1	0	0	1	0	0	1	1
0	1	0	1	1	0	1	1	0
0	1	1	1	1	0	1	1	0
1	0	0	0	0	0	0	1	1
1	0	0	0	1	0	0	1	1
1	0	0	1	1	0	1	1	0
1	0	1	1	1	0	1	1	0
1	1	0	0	0	0	1	1	0
					0	1	1	0

$$\left[\begin{array}{ccccc|ccc|c} 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & & \end{array}\right]$$

$$\overset{x}{U}\begin{bmatrix}0\\0\\1\end{bmatrix} + \overset{x}{W}\begin{bmatrix}h\end{bmatrix} + \vec{b} = \begin{bmatrix}h_{t+1}\end{bmatrix}$$

$$W$$

$$U\begin{bmatrix}0\\ \vdots\end{bmatrix} - \begin{bmatrix}x_1\\x_2\\ \vdots\end{bmatrix} + \begin{bmatrix}\ \end{bmatrix}\begin{bmatrix}0\\1\\ \vdots\end{bmatrix} + \begin{bmatrix}\ \end{bmatrix}$$

$$U \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \begin{bmatrix} x_3 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\overset{3\times 3}{U} \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} \overset{3\times 1}{x} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \overset{3\times 3}{W} \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} \overset{1\times 3}{h} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \overset{h}{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}$$

$$\downarrow$$

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{matrix} 1 \\ 1 \\ 1 \end{matrix}$$

$$\underline{e}$$

$$\downarrow$$

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{matrix} 0 \\ 2 \\ 1 \end{matrix}$$

$$x = (0, 0, 1) \begin{bmatrix} 1 & & -1 \\ & 1 & \\ & & 1 \end{bmatrix} = \begin{matrix} 0,0,1 \\ 0,1,1 \\ 1,0,1 \end{matrix}$$

0

3x3 3x1 3x1

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} \underline{0} & \underline{0} & \underline{0} \\ \ldots & \ldots & \ldots \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} \underline{1} & \underline{0} & \underline{0} \\ \underline{0} & \underline{1} & \underline{0} \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

$$Ux = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\} \qquad h = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

1   2   2   INP   Target

$$\begin{bmatrix} 0 \end{bmatrix} \quad \begin{bmatrix} \; \end{bmatrix} \quad \begin{bmatrix} 0 \end{bmatrix} \qquad \begin{bmatrix} 0 \end{bmatrix} \quad \begin{bmatrix} 0 \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}$$

$$\begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix} & & \end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix} - \begin{bmatrix}1\\0\end{bmatrix} - \begin{bmatrix}0\\1\end{bmatrix}$$

$$\begin{bmatrix} & & \end{bmatrix}\begin{bmatrix}0\\1\\1\end{bmatrix}^{2} = \begin{bmatrix}0\\0\\1\end{bmatrix}^{T}$$

1

$$\begin{bmatrix}0\\0\\0\end{bmatrix}^{0} = \begin{bmatrix}0\\0\\1\end{bmatrix}^{1} - \begin{bmatrix}0\\1\\0\end{bmatrix}^{I} = \begin{bmatrix}0\\-1\\1\end{bmatrix}^{T}$$

$$\begin{bmatrix}0\\0\\1\end{bmatrix}^{1} = \begin{bmatrix}0\\0\\1\end{bmatrix}^{1} - \begin{bmatrix}0\\1\\0\end{bmatrix}^{I} = \begin{bmatrix}0\\-1\\1\end{bmatrix}^{T}$$

$$\begin{bmatrix}1\\1\end{bmatrix}^{3} = \begin{bmatrix}0\\0\\1\end{bmatrix}^{\text{Target}}$$

$$W$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

2×3 $\quad 3\times 3$ $\quad 3\times 1$ $\quad \rightarrow$

$$U = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad W = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad V = \begin{bmatrix} 2 \\ -2 \end{bmatrix} \quad b_h = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$