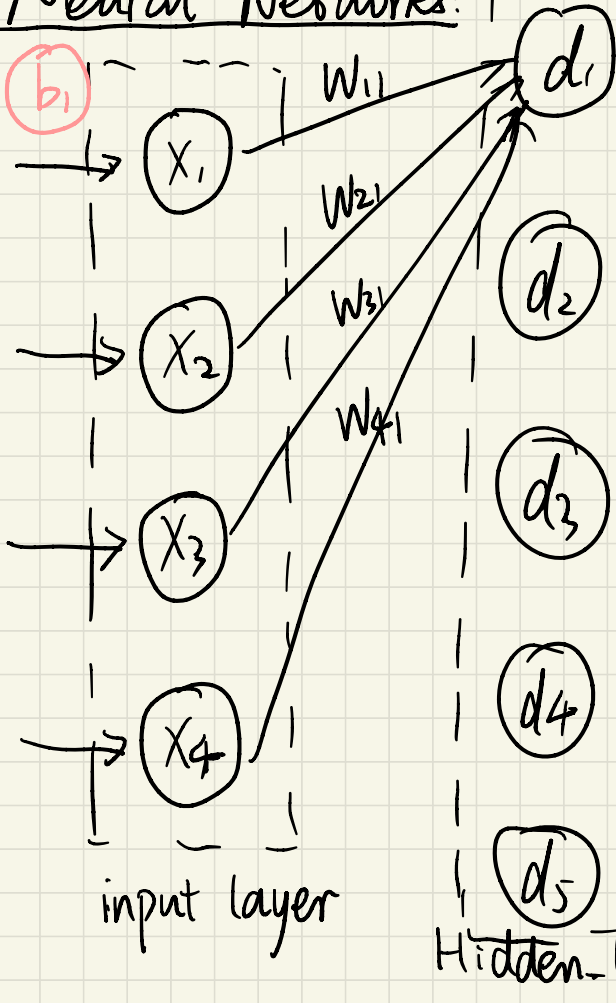# Neural Networks:

$W_1 = $ (transition matrix)

$W_1^T = (5 \times 4)$

$$W_1^T = \begin{bmatrix} W_{11} & W_{21} & W_{31} & W_{41} \\ W_{12} & W_{22} & W_{32} & W_{42} \\ W_{13} & W_{23} & W_{33} & W_{43} \\ W_{14} & W_{24} & W_{34} & W_{44} \\ W_{15} & W_{25} & W_{35} & W_{45} \end{bmatrix}$$

input layer

Hidden_layer_1

* W_{ij} is the weight connecting the node i in input layer and the node j in the hidden_layer.

batch size $= 32$ # of samples (inputs)

How do we define the shape of input?

→ [batch_size, d_0, d_1]

→ [batch_size, $(d_0,)$]

Activation Function

$$h_1 = g(a_1) = g(W_1^T X + b_1)$$

"activation function"

could be ReLu, tanh,

sigmoid and linear.

Why do we need activation function?

Ans: convert linear $\rightarrow$ nonlinear

Recall:
input layer

$X \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_5 \end{bmatrix}$ 🠐 not yet!

$(4 \times 1) \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix}$

(5×1)

$\underrightarrow{W^T X + b}$

linear
Transform
-ation

$\begin{bmatrix} W_{11}X_1 + W_{21}X_2 + W_{31}X_3 + W_{41}X_4 + b_1 \\ W_{12}X_1 + W_{22}X_2 + W_{32}X_3 + W_{42}X_4 + b_1 \\ W_{13}X_1 + W_{23}X_2 + W_{33}X_3 + \cdots + b_1 \\ W_{14}X_1 + W_{24}X_2 + W_{34}X_3 + \cdots + b_1 \\ W_{15}X_1 + \cdots \qquad\qquad + b_1 \end{bmatrix}$

$$\underset{\sigma.}{\underline{\text{Activation function}}} \longrightarrow \begin{bmatrix} \sigma(W_{11}X_1 + W_{21}X_2 + W_{31}X_3 + \cdots + b_2) \\ \sigma(W_{12}X_1 + W_{22}X_2 + \cdots + b_1) \\ \vdots \\ \sigma(W_{15}X_1 + W_{25}X_2 + \cdots + b_1). \end{bmatrix}$$

Now it is the $\downarrow$ hidden layer

DONE!

$$\begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \end{bmatrix}.$$

Model = Sequential ([

   ① Sequential is the function
we borrow from tf_keras to
build our model

   ② Dense → build a layer
   * input_shape = (1,)

   Why?

])

Summary : look at the

parameters :