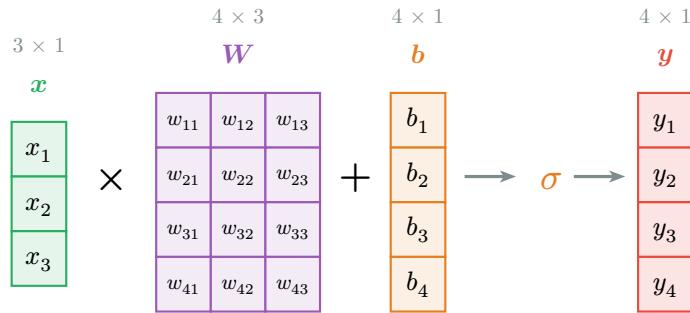


# Ogni Strato è un'operazione Matriciale



ogni riga di  $W$  contiene i pesi di un neurone

Uno strato = una moltiplicazione

$$y = \sigma(W \cdot x + b)$$

Dimensioni ( $n$  input,  $m$  neuroni):

$$x: n \times 1 \quad W: m \times n \quad b: m \times 1 \quad y: m \times 1$$

Rete completa = strati in sequenza:

$$y = \sigma_3(\sigma_2(\sigma_1(W_1 \cdot x + b_1) + b_2) + b_3)$$

Calcoli di uno strato in parallelo → GPU

# Ogni Strato è un'operazione Matriciale

$$\begin{array}{c} 3 \times 1 \\ \textcolor{green}{x} \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix} \end{array} \times \begin{array}{c} 4 \times 3 \\ \textcolor{violet}{W} \\ \begin{matrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \\ w_{41} & w_{42} & w_{43} \end{matrix} \end{array} + \begin{array}{c} 4 \times 1 \\ \textcolor{brown}{b} \\ \begin{matrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{matrix} \end{array} \rightarrow \sigma \rightarrow \begin{array}{c} 4 \times 1 \\ \textcolor{red}{y} \\ \begin{matrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{matrix} \end{array}$$

ogni riga di  $W$  contiene i pesi di un neurone

Ogni strato è una funzione:  $f_i : \mathbb{R}^{d_{i-1}} \rightarrow \mathbb{R}^{d_i}$

$$f_i(x) = \sigma_i(W_i \cdot x + b_i)$$

La rete è una composizione di funzioni:

$$F = f_n \circ f_{n-1} \circ \dots \circ f_1$$

$$\mathbf{y} = (f_n \circ f_{n-1} \circ \dots \circ f_1)(\mathbf{x})$$