

One Fully Connected Layer

- One fully-connected layer converts an input (column) vector x to an output (column) vector h :

$$h = f(Wx + b), \quad (1)$$

- W is a weight matrix of *input* columns and *output* rows,
- b a bias vector of length of *output*,
- $f(\cdot)$ is a non-linearity applied usually elementwise.

