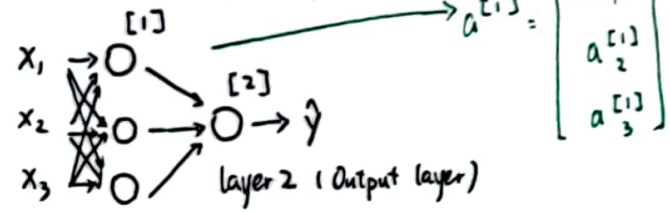
Neural Network



$$a^{[1]} = \begin{bmatrix} a_1^{[1]} \\ a_2^{[1]} \\ a_3^{[1]} \end{bmatrix}$$

$$z^{[1]} = W^{[1]} x + b^{[1]}$$

$$a^{[1]} = \sigma(z^{[1]})$$

$$z^{[2]} = W^{[2]} a^{[1]} + b^{[2]}$$

$$a^{[2]} = \sigma(z^{[2]})$$

$$\mathcal{L}(a^{[2]}, y)$$

Layer 1 Hidden Layer: $a^{[1]}$, $W^{[1]}$, $b^{[1]}$

Layer 2 output layer: $a^{[2]}$, $W^{[2]}$, $b^{[2]}$

假设我们有大量的 training example

$x \to a^{[2]} = \hat{y}$

$x^{(1)} \to a^{[2](1)} = \hat{y}^{(1)}$

$x^{(2)} \to a^{[2](2)} = \hat{y}^{(2)}$

$x^{(3)} \to a^{[2](3)} = \hat{y}^{(3)}$

$\vdots$

$x^{(m)} \to a^{[2](m)} = \hat{y}^{(m)}$

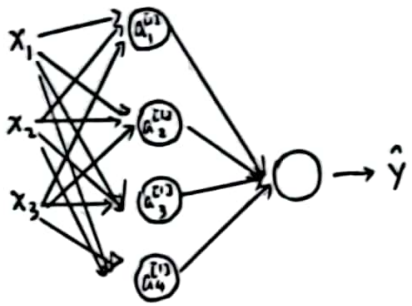
$a^{[2](i)}$ → example (i)

↳ layer

Code 如下：

```
for i to m
    z[1](i) = W[1] x(i) + b[1]
    a[1](i) = σ(z[1](i))
    z[2](i) = W[2] a[1](i) + b[2]
    a[2](i) = σ(z[2](i))
```

Neural Network 的表示方法



$$\begin{cases} z_1^{[1]} = (w_1^{[1]})^T \cdot x + b_1^{[1]} \\ a_1^{[1]} = \sigma(z_1^{[1]}) \end{cases}$$

$$\begin{cases} z_2^{[1]} = (w_2^{[1]})^T \cdot x + b_2^{[1]} \\ a_2^{[1]} = \sigma(z_2^{[1]}) \end{cases}$$

$$\begin{cases} z_3^{[1]} = (w_3^{[1]})^T \cdot x + b_3^{[1]} \\ a_3^{[1]} = \sigma(z_3^{[1]}) \end{cases}$$

$$\begin{cases} z_4^{[1]} = (w_4^{[1]})^T \cdot x + b_4^{[1]} \\ a_4^{[1]} = \sigma(z_4^{[1]}) \end{cases}$$

Vector 的表示结构

$$z^{[1]} = \begin{bmatrix} -(w_1^{[1]})^T- \\ -(w_2^{[1]})^T- \\ -(w_3^{[1]})^T- \\ -(w_4^{[1]})^T- \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} b_1^{[1]} \\ b_2^{[1]} \\ b_3^{[1]} \\ b_4^{[1]} \end{bmatrix} = \begin{bmatrix} (w_1^{[1]})^T x_1 + b_1^{[1]} \\ (w_2^{[1]})^T x_2 + b_2^{[1]} \\ (w_3^{[1]})^T x + b_3^{[1]} \\ (w_4^{[1]})^T x + b_4^{[1]} \end{bmatrix} = \begin{bmatrix} z_1^{[1]} \\ z_2^{[1]} \\ z_3^{[1]} \\ z_4^{[1]} \end{bmatrix}$$

$W^{[1]}$ (4,3) $\times$ (3,1) $+$ $b^{[1]}$ (4,1)

$$a^{[1]} = \begin{bmatrix} a_1^{[1]} \\ a_2^{[1]} \\ a_3^{[1]} \\ a_4^{[1]} \end{bmatrix}$$

$$z^{[1]} = W^{[1]} x + b^{[1]}$$

$$a^{[1]} = \sigma(z^{[1]})$$

$$z^{[2]} = W^{[2]} x + b^{[2]}$$

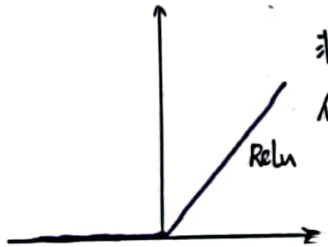
$$a^{[2]} = \sigma(z^{[2]})$$

<第一层线性> 输入特征向量

$z^{[1]} = W^{[1]} x + b^{[1]}$

↓ 线性输出　权重矩阵 (weight matrix)　(bias vector) 偏置向量

$a^{[1]} = \sigma(z^{[1]})$ ↳ 非线性　σ 是激活函数 (sigmoid / ReLu)



非线性帮助神经网络作出多样的复杂的决策

权重矩阵，weight matrix？

作用：调整输入信号的强度

权重越大，该输入特征对输出影响越大；权重越小，该输入特征对输出影响越小

偏置矩阵，Bias vector？

作用：通过平移激活函数来调整非0的输出（神经元）