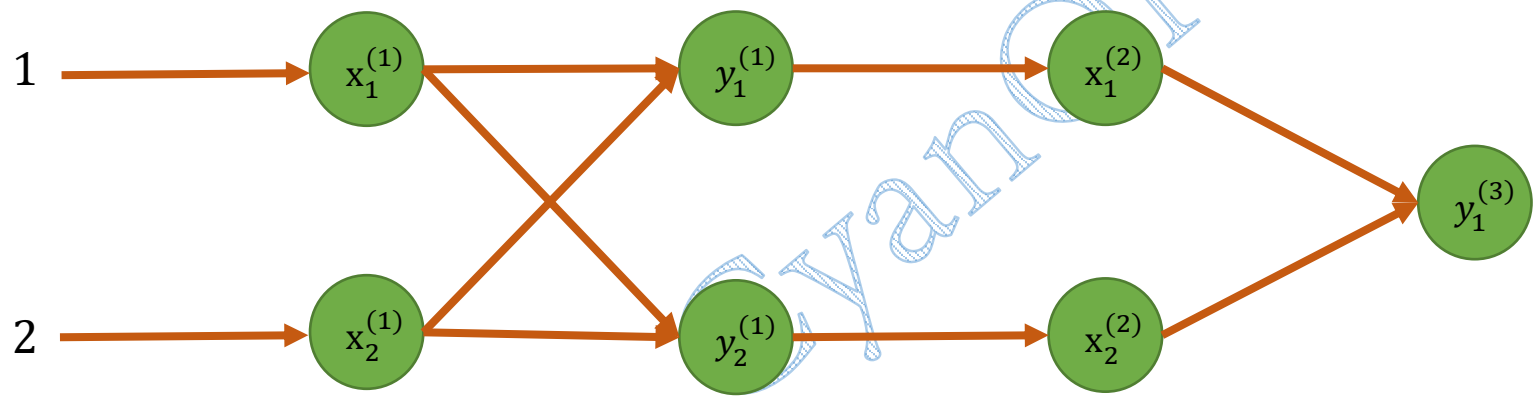
Input: $x_1^{(1)} = 1 \quad x_2^{(1)} = 2$

Lable: $\dot{A} = 2$

激活函数 $y = ReLU(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$

Learning rate(Lr) =0.01

损失函数 最小均方误差-Minimum Mean Squared Error (MMSE) $J(w, b) = \text{Loss} = \frac{1}{2}(A - \dot{A})^2$



初始化权重：

$$\boldsymbol{w_{1,1}^{(1)} = w_{1,2}^{(1)} = w_{2,1}^{(1)} = w_{1,1}^{(2)} = 1}$$

$$\boldsymbol{w_{2,2}^{(1)} = w_{2,1}^{(2)} = -1}$$

$$\boldsymbol{\theta_{1,1}^{(1)} = \theta_{1,2}^{(1)} = \theta_{1,1}^{(2)}}=1$$

则：

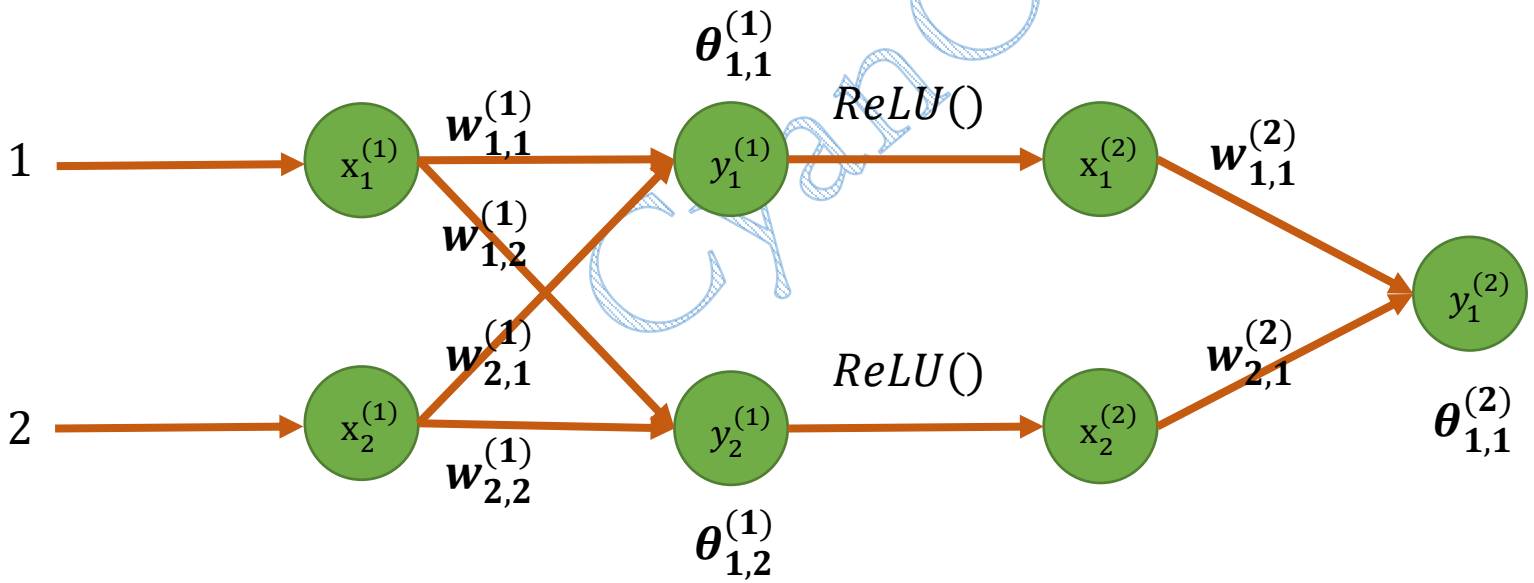
$$y_1^{(1)} = x_1^{(1)} * w_{1,1}^{(1)} + x_2^{(1)} * w_{2,1}^{(1)} + \theta_{1,1}^{(1)} = 4$$

$$x_1^{(2)} = ReLU\left(y_1^{(1)}\right) = 4$$

$$y_2^{(1)} = x_1^{(1)} * w_{1,2}^{(1)} + x_2^{(1)} * w_{2,2}^{(1)} + \theta_{1,2}^{(1)} = 0$$

$$x_2^{(2)} = ReLU\left(y_2^{(1)}\right) = 0$$

$$\mathrm{A} = y_1^{(2)} = x_1^{(2)} * w_{1,1}^{(2)} + x_2^{(2)} * w_{2,1}^{(2)} + \theta_{1,1}^{(2)} = 5$$



$w^{(1)}_{1,1_grad}$表示参数$w^{(1)}_{1,1}$的梯度

$$\mathrm{J(w,b)} = \mathrm{Loss} = \frac{1}{2}(A - \dot{A})^2 = 4.5$$

$$\widetilde{w}^{(1)}_{1,1} = w^{(1)}_{1,1} - \mathrm{Lr} * w^{(1)}_{1,1\_grad}$$

$$w^{(1)}_{1,1\_grad} = \frac{\partial J}{\partial w^{(1)}_{1,1}} = \frac{\partial J}{\partial A} * \frac{\partial A}{\partial x^{(2)}_1} * \frac{\partial x^{(2)}_1}{\partial y^{(1)}_1} * \frac{\partial y^{(1)}_1}{\partial w^{(1)}_{1,1}} = \left(A - \dot{A}\right) * w^{(2)}_{1,1} * 1 * x^{(1)}_1 = 3$$

$$\widetilde{w}^{(1)}_{1,1} = w^{(1)}_{1,1} - \mathrm{Lr} * w^{(1)}_{1,1_{grad}} = 1 - 0.01 * 3 = 0.97$$

同理可求得$\frac{\partial J}{\partial w^{(1)}_{1,2}}$, $\frac{\partial J}{\partial w^{(1)}_{2,1}}$, $\frac{\partial J}{\partial w^{(1)}_{2,2}}$, $\frac{\partial J}{\partial w^{(2)}_{1,1}}$, $\frac{\partial J}{\partial w^{(2)}_{2,1}}$, $\frac{\partial J}{\partial \theta^{(1)}_{1,1}}$, $\frac{\partial J}{\partial \theta^{(1)}_{1,2}}$, $\frac{\partial J}{\partial \theta^{(2)}_{1,1}}$