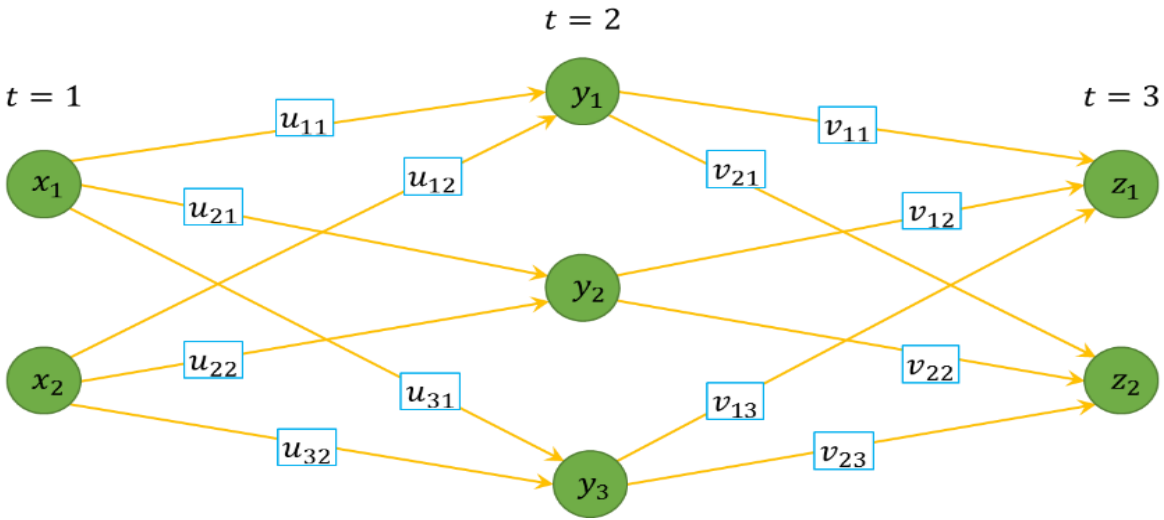
参考答案：作业1推导

设有三层前馈神经网络如下图所示，其中t表示网络的层数，u_{ij}和v_{ij}分别为各层的参数，用E_k表示样本x_k产生的误差，激活函数$S(x)=\frac{1}{1+exp\ (-x)}$，$f_1(x_k)$和$f_2(x_k)$表示两个输出节点的目标值，第2、3层节点内的值代表通过激活函数后的结果。请推导参数的梯度$\frac{\partial E_k}{\partial v_{11}}$、$\frac{\partial E_k}{\partial v_{21}}$和$\frac{\partial E_k}{\partial u_{11}}$。



参考答案：作业1推导

前馈神经网络的目标函数为最小化平方误差，又有$S'(x) = S(x)\big(1 - S(x)\big)$，可求：

$$\frac{\partial E_k}{\partial v_{11}} = \frac{\partial E_k}{\partial z_1}\frac{\partial z_1}{\partial v_{11}} = -(f_1(x_k) - z_1)z_1(1 - z_1)y_1$$

$$E_k = \frac{1}{2}\sum_{j=1}^{l}(z_k - f_k(x_k))^2 \qquad z_1 = S(\frac{y_1v_{11} + y_2v_{21} + y_3v_{31}}{m})$$

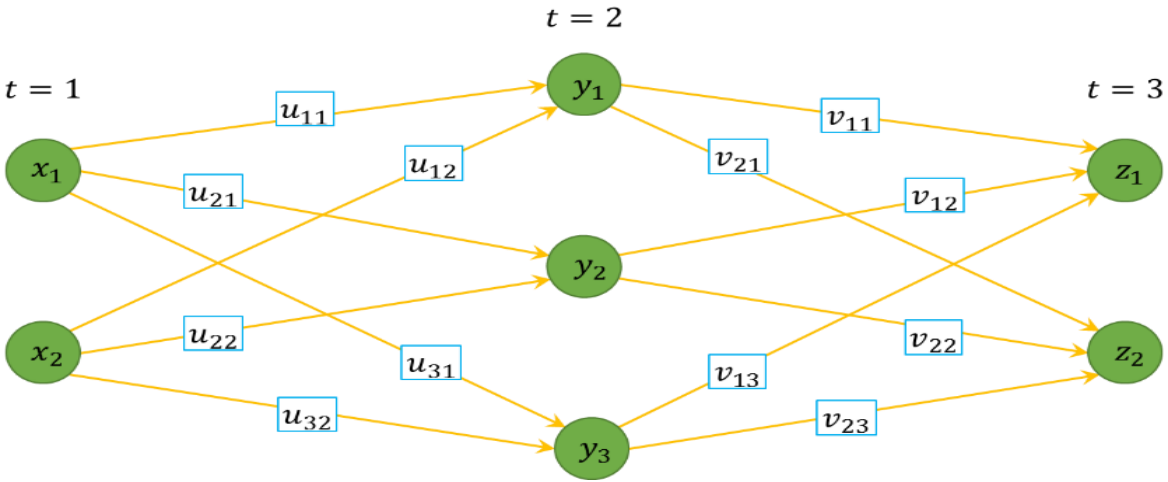
$$\begin{aligned}\frac{\partial E_k}{\partial v_{11}} = \frac{\partial E_k}{\partial z_1}\frac{\partial z_1}{\partial v_{11}} &= \big(z_1 - f_1(x_k)\big)S'(m)y_1 \\ &= \big(z_1 - f_1(x_k)\big)S(m)(1 - S(m))y_1 \\ &= -(f_1(x_k) - z_1)z_1(1 - z_1)y_1\end{aligned}$$

参考答案：作业1推导

前馈神经网络的目标函数为最小化平方误差，又有$S'(x) = S(x)\big(1 - S(x)\big)$，可求：

同理可得

$$\frac{\partial E_k}{\partial v_{21}} = \frac{\partial E_k}{\partial z_2}\frac{\partial z_2}{\partial v_{21}} = -(f_2(x_k) - z_2)z_2(1 - z_2)y_1$$



参考答案：作业1推导

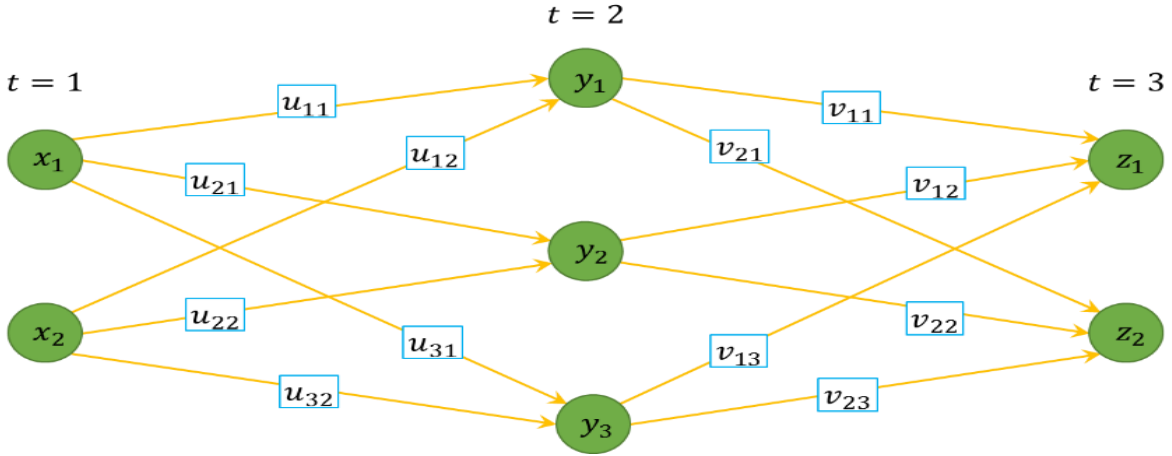
$$\frac{\partial E_k}{\partial u_{11}} = \frac{\partial E_k}{\partial z_1}\frac{\partial z_1}{\partial y_1}\frac{\partial y_1}{\partial u_{11}} + \frac{\partial E_k}{\partial z_2}\frac{\partial z_2}{\partial y_1}\frac{\partial y_1}{\partial u_{11}}$$

$$E_k = \frac{1}{2}\sum_{j=1}^{l}(z_k - f_k(x_k))^2$$

$$z_1 = S(\frac{y_1 v_{11} + y_2 v_{12} + y_3 v_{13}}{m})$$

$$y_1 = S(x_1 u_{11} + x_2 u_{12}) = S(p)$$

$$z_2 = S(\frac{y_1 v_{21} + y_2 v_{22} + y_3 v_{23}}{n})$$



参考答案：作业1推导

$$\frac{\partial E_k}{\partial u_{11}} = \frac{\partial E_k}{\partial z_1}\frac{\partial z_1}{\partial y_1}\frac{\partial y_1}{\partial u_{11}} + \frac{\partial E_k}{\partial z_2}\frac{\partial z_2}{\partial y_1}\frac{\partial y_1}{\partial u_{11}}$$

$$E_k = \frac{1}{2}\sum_{j=1}^{l}(z_k - f_k(x_k))^2 \qquad z_1 = S(\frac{y_1 v_{11} + y_2 v_{12} + y_3 v_{13}}{m})$$

$$y_1 = S(x_1 u_{11} + x_2 u_{12}) = S(p)$$

$$\frac{\partial E_k}{\partial z_1}\frac{\partial z_1}{\partial y_1}\frac{\partial y_1}{\partial u_{11}} = (z_1 - f_1(x_k))S'(m)v_{11}S'(p)x_1$$

$$= (z_1 - f_1(x_k))S(m)(1 - S(m)v_{11}S(p)(1 - S(p)x_1$$

$$= -(f_1(x_k) - z_1)z_1(1 - z_1)v_{11}y_1(1 - y_1)x_1$$

参考答案：作业1推导

$$E_k = \frac{1}{2}\sum_{j=1}^{l}(z_k - f_k(x_k))^2$$

$$z_2 = S(\frac{y_1 v_{21} + y_2 v_{22} + y_3 v_{23}}{n})$$

$$y_1 = S(x_1 u_{11} + x_2 u_{12}) = S(p)$$

同理可得

$$\frac{\partial E_k}{\partial z_2}\frac{\partial z_2}{\partial y_1}\frac{\partial y_1}{\partial u_{11}} = (z_2 - f_2(x_k))S'(n)v_{21}S'(p)x_1$$

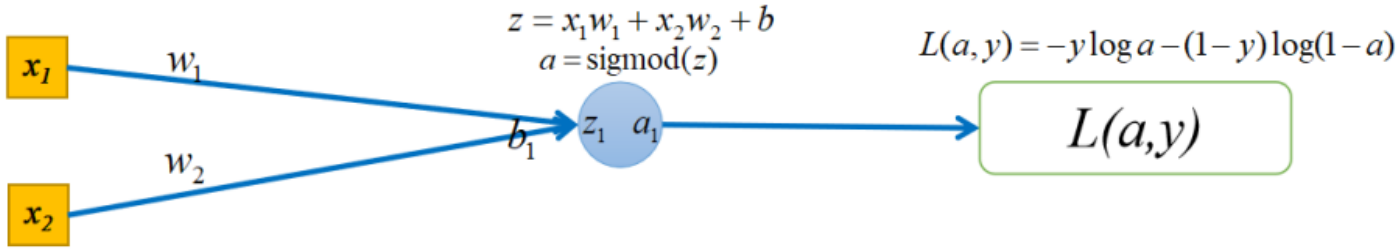
$$= -(f_2(x_k) - z_2)z_2(1 - z_2)v_{21}y_1(1 - y_1)x_1$$

$$\frac{\partial E_k}{\partial u_{11}} = \frac{\partial E_k}{\partial z_1}\frac{\partial z_1}{\partial y_1}\frac{\partial y_1}{\partial u_{11}} + \frac{\partial E_k}{\partial z_2}\frac{\partial z_2}{\partial y_1}\frac{\partial y_1}{\partial u_{11}}$$

$$= -(f_1(x_k) - z_1)z_1(1 - z_1)v_{11}y_1(1 - y_1)x_1 - (f_2(x_k) - z_2)z_2(1 - z_2)v_{21}y_1(1 - y_1)x_1$$

参考答案：作业2计算

假设最开始初始化$w_1 = 0.3$，$w_2 = 0.4$，$b = 1$，本轮训练样本为[(2,3),0]，损失函数选用的对数损失，学习率$\eta = 0.1$，求第一次和第二次学习得到的输出节点值a（写出计算公式和计算过程）。



参考答案：作业2计算

(1) 第一次训练的正向传播过程如下：

$$z = x_1 w_1 + x_2 w_2 + b = 2 \times 0.3 + 3 \times 0.4 + 1 = 2.8$$

$$a = sigmoid(z) = sigmoid(2.8) = 0.943$$

$$L(a, y) = -y\log a - (1 - y)\log(1 - a) = -\log(1 - 0.943) = 2.859$$

参考答案：作业2计算

(2) 第一次训练的反向传播过程如下：

$$\frac{\partial L}{\partial w_1}=\frac{\partial L}{\partial a}\frac{\partial a}{\partial z}\frac{\partial z}{\partial w_1}=\left(-\frac{y}{a}+\frac{1-y}{1-a}\right)\cdot a(1-a)\cdot x_1=1.885$$

$$\frac{\partial L}{\partial w_2}=\frac{\partial L}{\partial a}\frac{\partial a}{\partial z}\frac{\partial z}{\partial w_2}=\left(-\frac{y}{a}+\frac{1-y}{1-a}\right)\cdot a(1-a)\cdot x_2=2.828$$

$$\frac{\partial L}{\partial b}=\frac{\partial L}{\partial a}\frac{\partial a}{\partial z}\frac{\partial z}{\partial b}=\left(-\frac{y}{a}+\frac{1-y}{1-a}\right)\cdot a(1-a)=0.943$$

更新参数：

$$w_1=w_1-\eta\frac{\partial L}{\partial w_1}=0.111$$

$$w_2=w_2-\eta\frac{\partial L}{\partial w_2}=0.117$$

$$b=b-\eta\frac{\partial L}{\partial b}=0.906$$

参考答案：作业2计算

(3) 第二次训练的正向传播过程如下：

$$z = x_1 w_1 + x_2 w_2 + b = 2 \times 0.111 + 3 \times 0.117 + 0.906 = 1.479$$

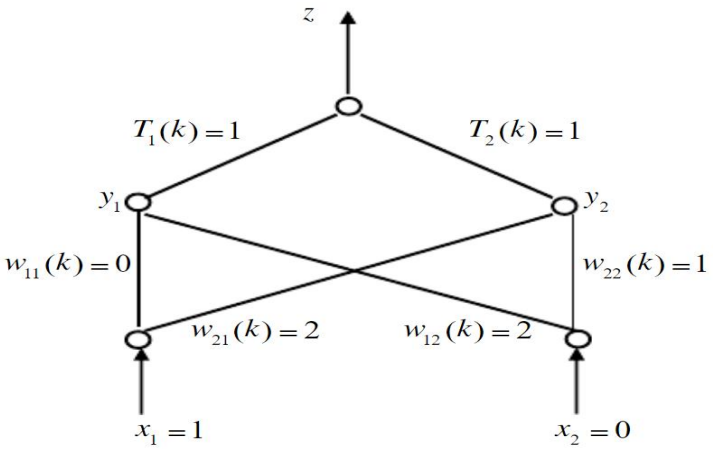
$$a = sigmoid(z) = sigmoid(1.479) = 0.814$$

参考答案：作业3计算

对如下的BP神经网络，学习率$\eta = 1$，各点的偏差$\theta = 0$。激活函数为：

$$f(x) = \begin{cases} x & x \geq 1 \\ 1 & x < 1 \end{cases}$$

输入样本$x_1 = 1, x_2 = 0$，输出节点z的期望输出为$z' = 1$，对于第k次学习得到的权值分别为$w_{11}(k) = 0, w_{12}(k) = 2, w_{21}(k) = 2, w_{22}(k) = 1, T_1(k) = 1, T_2(k) = 1$，损失函数是均方误差，求第$k$次和$k+1$次学习得到的输出节点值$z(k)$和$z(k+1)$（写出计算公式和计算过程）。



参考答案：作业3计算

(1) 第 k 次训练的正向传播过程如下：

$$y_1 = f\left(\sum_{j=1}^{2} w_{1j}\, x_j + \theta\right) = f(0 \times 1 + 2 \times 0) = f(0) = 1$$

$$y_2 = f\left(\sum_{j=1}^{2} w_{2j}\, x_j + \theta\right) = f(2 \times 1 + 1 \times 0) = f(2) = 2$$

$$z = f\left(\sum_{i=0}^{2} T_i y_i + \theta\right) = f(1 \times 1 + 1 \times 2) = f(3) = 3$$

$$L = \frac{1}{2}(z - z')^2 = \frac{1}{2}(3 - 1)^2 = 2$$

参考答案：作业3计算

(2) 第 k 次训练的反向传播过程如下：

$$\frac{\partial L}{\partial T_1} = \frac{\partial L}{\partial z}\frac{\partial z}{\partial T_1} = (z - z') \cdot (1 \cdot y_1) = (3 - 1) \times (1 \times 1) = 2$$

$$\frac{\partial L}{\partial T_2} = \frac{\partial L}{\partial z}\frac{\partial z}{\partial T_2} = (z - z') \cdot (1 \cdot y_2) = (3 - 1) \times (1 \times 2) = 4$$

$$\frac{\partial L}{\partial w_{11}} = \frac{\partial L}{\partial z}\frac{\partial z}{\partial y_1}\frac{\partial y_1}{\partial w_{11}} = (z - z') \cdot (1 \cdot T_1) \cdot (0 \cdot x_1) = 0$$

$$\frac{\partial L}{\partial w_{12}} = \frac{\partial L}{\partial z}\frac{\partial z}{\partial y_1}\frac{\partial y_1}{\partial w_{12}} = (z - z') \cdot (1 \cdot T_1) \cdot (0 \cdot x_2) = 0$$

$$\frac{\partial L}{\partial w_{21}} = \frac{\partial L}{\partial z}\frac{\partial z}{\partial y_2}\frac{\partial y_2}{\partial w_{21}} = (z - z') \cdot (1 \cdot T_2) \cdot (1 \cdot x_1) = (3 - 1) \times (1 \times 1) \times (1 \times 1) = 2$$

$$\frac{\partial L}{\partial w_{22}} = \frac{\partial L}{\partial z}\frac{\partial z}{\partial y_2}\frac{\partial y_2}{\partial w_{22}} = (z - z') \cdot (1 \cdot T_2) \cdot (1 \cdot x_2) = (3 - 1) \times (1 \times 1) \times (1 \times 0) = 0$$

参考答案：作业3计算

更新参数：

$$T_1(k+1) = T_1(k) - \eta \frac{\partial L}{\partial T_1} = 1 - 1 \times 2 = -1$$

$$T_2(k+1) = T_2(k) - \eta \frac{\partial L}{\partial T_2} = 1 - 1 \times 4 = -3$$

$$w_{11}(k+1) = w_{11}(k) - \eta \frac{\partial L}{\partial w_{11}} = 0 - 1 \times 0 = 0$$

$$w_{12}(k+1) = w_{12}(k) - \eta \frac{\partial L}{\partial w_{12}} = 2 - 1 \times 0 = 2$$

$$w_{21}(k+1) = w_{21}(k) - \eta \frac{\partial L}{\partial w_{21}} = 2 - 1 \times 2 = 0$$

$$w_{22}(k+1) = w_{22}(k) - \eta \frac{\partial L}{\partial w_{22}} = 1 - 1 \times 0 = 1$$

参考答案：作业3计算

(3) 第 k+1 次训练的正向传播过程如下：

$$y_1 = f\left(\sum_{j=1}^{2} w_{1j}\, x_j + \theta\right) = f(0 \times 1 + 2 \times 0) = f(0) = 1$$

$$y_2 = f\left(\sum_{j=1}^{2} w_{2j}\, x_j + \theta\right) = f(0 \times 1 + 1 \times 0) = f(0) = 1$$

$$z = f\left(\sum_{i=0}^{2} T_i y_i + \theta\right) = f(-1 \times 1 - 3 \times 1) = f(-4) = 1$$

$$L = \frac{1}{2}(z - z')^2 = \frac{1}{2}(1 - 1)^2 = 0$$