HW 6

1、设 $A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} = [1 \ 0 \ 0]_{4\times 3}$

$A_2 = [0 \ 1 \ 0]_{2\times 3}$ $\qquad A_3 = [0 \ 0 \ 1]_{1\times 3}$

$\therefore \ h_1 = \cos(W_1 x + A_1 \vec{b})$

$h_2 = \cos(W_2 h_1 + A_2 \vec{b})$，其中初始值 $W_2 = W_1^T$

$y = \mathrm{sigmoid}(W_3 h_2 + A_3 \vec{b})$

代入初始值计算：$h_1 = [0.863,\ 0.853,\ 0.863,\ 0.863]^T$

$h_2 = [0.452,\ 0.453,\ 0.452]^T$

$y = 0.63]$

(2) 设 $E = \frac{1}{2} \|\hat{y} - y\|_2^2$

$$\frac{\partial E}{\partial W_3} = \frac{\partial E}{\partial \hat{y}} \quad \frac{\partial \hat{y}}{\partial o} \quad \frac{\partial o}{\partial W_3} = (\hat{y} - y) \cdot \hat{y}(1-\hat{y}) \cdot h_2^T$$

$$= [[-0.032],\ -0.0328,\ -0.032]]$$

(3) $w_3^{(1)} = w_3^{(0)} - 2 \cdot \frac{\partial E}{\partial w_3}$

$= [0.2033, \quad 0.2033, \quad 0.4033]$

3、(1) Input: 3

output: 4

kernel width: 5或6 （$2\times48-1=95$， $100-95=5$ 故可能是5或6）

(2) 4个卷积核，因为不妨考虑中间9个格子

1	2	3
4	5	6
7	8	9

3×3卷积核不妨设为

0	0	0
0	1	0
0	0	0

按顺序从1开始时，1、3、7、9均可保留

2开始时，2、8保留；4开始则4、6保留，5开始则5保留，至此1-9信息全部被留下，故至少4个核

注：另一个角度分析，假设要保存$N\times N$的特征，按当前核一次卷积可得$\frac{1}{4}N\times N$的特征（不计边界）

故至少要 $N\times N / (\frac{1}{4} N\times N) = 4$ 个卷积核

4、设输入为 $X = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 2 \end{bmatrix}$ $Y = \begin{bmatrix} 1 & 1.5 \\ 1 & 1.5 \end{bmatrix}$ （网络输出值）

$$D = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$$ （理论输出值） $$Y - D = \begin{bmatrix} 0 & 0.5 \\ -1 & 0.5 \end{bmatrix}$$

$$E = \sum_{i,j} \frac{1}{2}(y_{ij} - d_{ij})^2$$

$$\frac{\partial E}{\partial H_{m11}} = \sum_{i,j} (y_{ij} - d_{ij}) \cdot \frac{\partial y_{ij}}{\partial H_{m11}}$$

$$= 1 \times 0 + 1 \times 0.5 + 1 \times (-1) + 1 \times 0.5 = 0$$

$$\frac{\partial E}{\partial H_{m12}} = 1 \times 0 + 2 \times 0.5 + 1 \times (-1) + 2 \times 0.5 = 1$$

$$\frac{\partial E}{\partial H_{m21}} = 1 \times 0 + 1 \times 0.5 - 1 \times 1 + 1 \times 0.5 = 0$$

$$\frac{\partial E}{\partial H_{m22}} = 1 \times 0 + 2 \times 0.5 + 1 \times (-1) + 2 \times 0.5 = 1$$

$$\therefore \Delta H_m = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

$$\therefore H_m^{(1)} = H_m^{(0)} - \alpha \Delta H_m = \begin{bmatrix} 0.2 & -0.8 \\ 0.3 & -0.7 \end{bmatrix}$$