problem 01

(1) for example X_i, $net_{ik}=\omega_{ik}X_i+b_{ij}$, $z_{ik}=\frac{e^{net_{ik}}}{\sum_{k=1}^c e^{net_{ik}}}$.

the lost function: $L_i=-\sum_{k=1}^c[y_{ik}log(z_{ik})]$

the cost function: $C=\frac{1}{N}\sum_{i=1}^NL_i=-\frac{1}
{N}\sum_{i=1}^N\sum_{k=1}^c[y_{ik}log(z_{ik})]$

the gradient of hidden-to-output weights: $\frac{\partial C}{\partial net_{ik}}=\frac{\partial C}{\partial z_{ij}}\frac{\partial z_{ij}}{\partial net_{ik}}$

in this formula, $\frac{\partial C}{\partial z_{ij}}=-\frac{1}{N}\sum_{i=1}^N\sum_{j=1}^c[\frac{y_{ij}}{z_{ij}}]$

when $j=k$: $\frac{\partial z_{ik}}{\partial net_{ik}}=\frac{\partial}{\partial net_{ik}}\frac{e^{net_{ik}}}{\sum_{k=1}^c e^{net_{ik}}}=\frac{e^{net_{ik}}\sum_{k=1}^c e^{net_{ik}}-e^{2net_{ik}}}{(\sum_{k=1}^c e^{net_{ik}})^2}=z_{ik}-z_{ik}^2$

when $j\neq k$: $\frac{\partial z_{ij}}{\partial net_{ik}}=\frac{\partial}{\partial net_{ik}}\frac{e^{net_{ij}}}{\sum_{k=1}^c e^{net_{ik}}}=-\frac{e^{net_{ij}}e^{net_{ik}}}{(\sum_{k=1}^c e^{net_{ik}})^2}=-z_{ij}z_{ik}$

take them into the original fomula,
$\frac{\partial C}{\partial net_{ik}}=-\frac{1}{N}\sum_{i=1}^N(\sum_{j\neq k}^c\frac{y_{ij}}{z_{ij}}(-z_{ij}z_{ik})+\frac{y_{ik}}{z_{ik}}(z_{ik}-z_{ik}^2))=-\frac{1}{N}\sum_{i=1}^N(\sum_{j=1}^c(-y_{ij}z_{ik})+y_{ik})$

(2)don't really understand

problem 02

(1)

$h_{11}=f_{11}(x_1)=\left\{\begin{matrix} -1, & x_1\leq 0.5 \\ 1, & x_1>0.5 \end{matrix}\right.$

$h_{12}=f_{12}(x_2)=\left\{\begin{matrix} -1, & x_2\leq 0.5 \\ 1, & x_2>0.5 \end{matrix}\right.$

$g=h_{11}*h_{12}$

(2)

$h_{21}=f_{21}(x_1)=\left\{\begin{matrix} 1 , & 1<x_1\leq 1.5 \\ -1 , & x_1>1.5
\end{matrix}\right.$

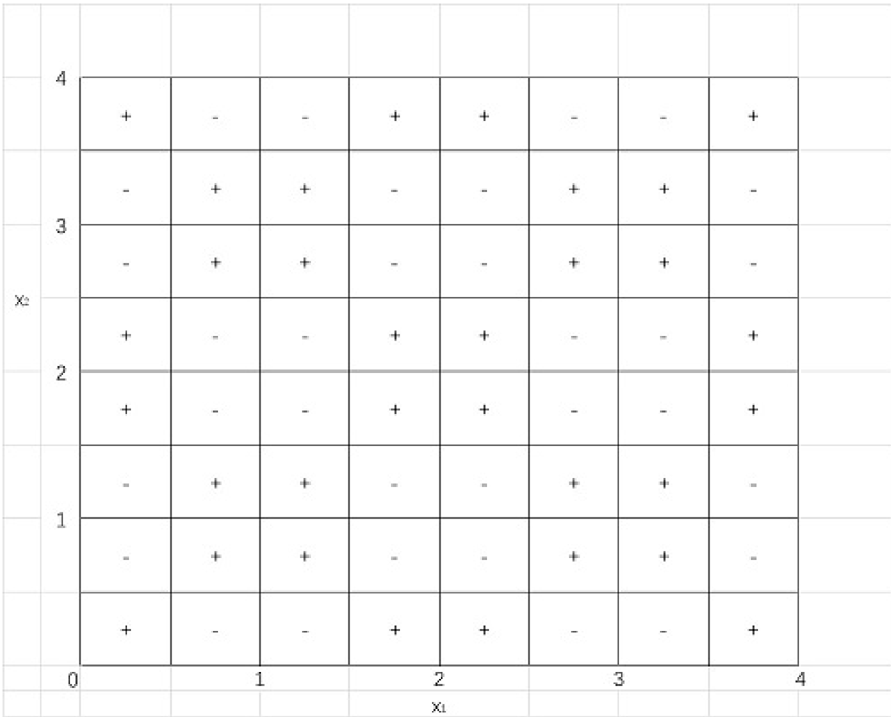
$=-f_{11}(x_1-1)$

$$h_{22}=f_{22}(x_2)=\left\{\begin{matrix} 1 , & 1<x_2\leq 1.5 \\ -1 , & x_2>1.5 \end{matrix}\right.$$

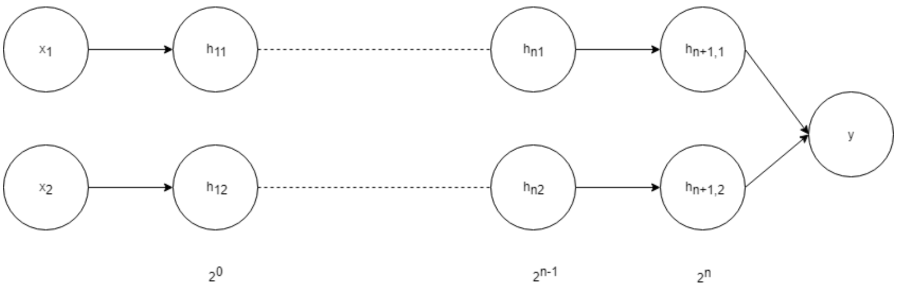
$= -f_{12}(x_2 - 1)$

$g = h_{21} * h_{22}$

(3)data are mirror symmetry to the 2^n axis.



(4)

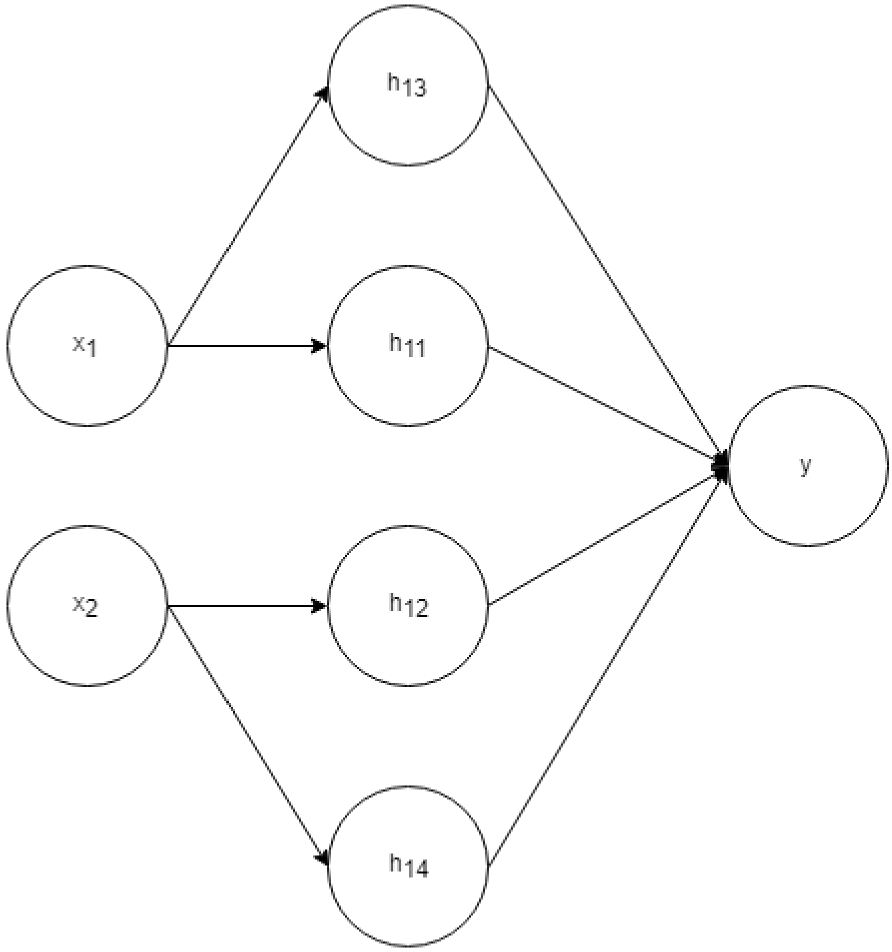


when $n \geq 1$:

$h_{n1} = f_{n1}(x_1) = -f_{n-1}(x_1 - 2^{n-1})$

$h_{n2} = f_{n2}(x_2) = -f_{n-1}(x_2 - 2^{n-1})$

(5)



$h_{11} = (-1)^{x_1|0.5}$

$h_{13} = (-1)^{x_1|1}$

$h_{12} = (-1)^{x_2|0.5+1}$

$h_{14} = (-1)^{x_2|1+1}$

$y = h_{11} * h_{12} * h_{13} * h_{14}$