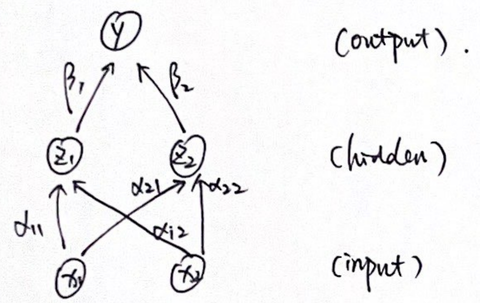
机器学习与人工智能.　　Deep Learning.

1. Neural Network.

It is like a combination of logistic regression.

But not really. ∵ no target for hidden units.

例#1. NN w/1 Hidden Layers and 2 Hidden Units.



激活函数. (Nonlinear) → 增加Net的非线性 (即拟合能力)

σ: activation function.

if σ is sigmoid $\sigma(a) = \frac{1}{1+\exp(-a)}$.

$x_i \in \mathbb{R}$.

$z_i \in (0,1)$ if σ is sigmoid.

$z_i \in \mathbb{R}$ more generally.

$z_1 = \sigma(\alpha_{11} x_1 + \alpha_{12} x_2 + \alpha_{10})$.

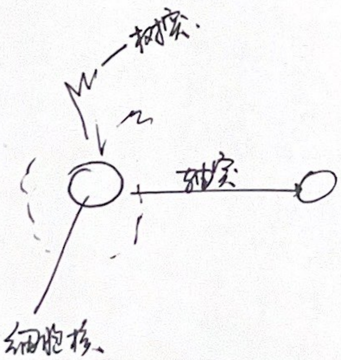
$z_2 = \sigma(\alpha_{21} x_1 + \alpha_{22} x_2 + \alpha_{20})$

$y = \sigma(\beta_1 z_1 + \beta_2 z_2 + \beta_0)$.

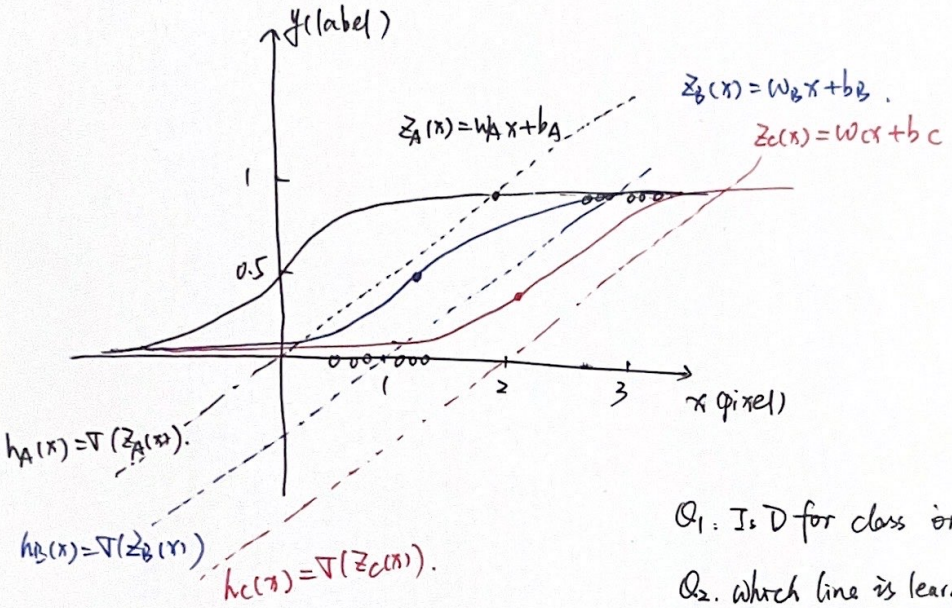
$= \sigma(\beta_1 \sigma(\alpha_{11} x_1 + \alpha_{12} x_2 + \underline{\underline{\alpha_{10}}})$

$+ \beta_2 \sigma(\alpha_{21} x_1 + \alpha_{22} x_2 + \underline{\underline{\alpha_{20}}})$

$+ \underline{\underline{\beta_0}})$. Don't forget intercepts.



1D Face Recognition

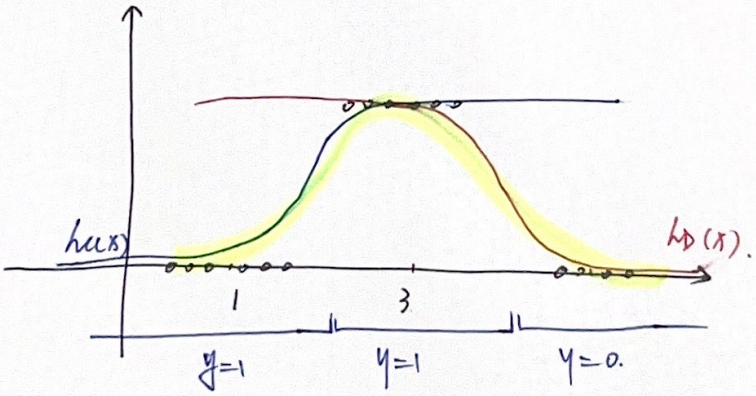


Q_1: Is D for class or reg? both

Q_2. which line is learned by linear Reg? $z_B(x)$

Q_3 which sigmoid is learned by logistic Reg? $h_C(x)$

Q_4 what is the D.B. for $h_C(x)$ point $x=2$



$$h_E(x) = \sigma\left(\frac{h_C(x) + h_D(x)}{2}\right)$$

D.B. is a nonlinear fnc. of x.

3. Backpropogation 误差反向传播

→ Not a NN diagram

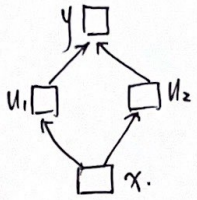
Computation graphs

Chain Rule.

Def #1. $y = f(u)$, $u = g(x)$

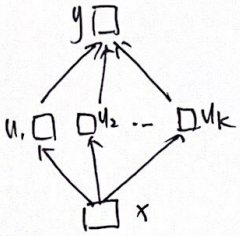
$$\frac{\partial y}{\partial x} = \frac{\partial y}{\partial u}\frac{\partial u}{\partial x}.$$

Def #2. $y = f(u_1, u_2)$, $u_2 = g_2(x)$, $u_1 = g_1(x)$



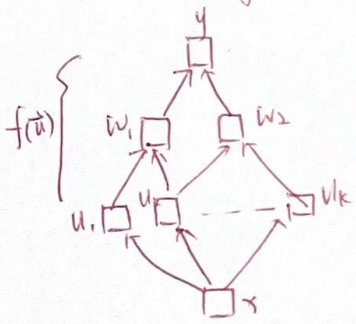
$$\frac{\partial y}{\partial x} = \frac{\partial y}{\partial u_1}\frac{\partial u_1}{\partial x} + \frac{\partial y}{\partial u_2}\cdot\frac{\partial u_2}{\partial x}.$$

Def #3. $y = f(\vec{u})$, $\vec{u} = g(x)$.



$$\frac{\partial y}{\partial x} = \sum_{k=1}^{K}\frac{\partial y}{\partial u_k}\frac{\partial u_k}{\partial x}.$$

★ holds for any intermediate quantities $\vec{u}$



Derivative of a Sigmoid.

$$S = \frac{1}{1+\exp(-b)}. \qquad \frac{ds}{db} = S(1-S).$$