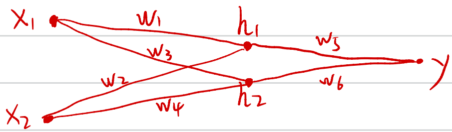

Q1: 为什么要求梯度？

Q2: 求谁的梯度？

$y = f_w(x)$ 　　w为学习参数

目标: $E = \min_{w} \sum_{x} \| f_w(x) - \hat{y} \|^2$

求得一组w使得E最小.

方法: 随机初始化w, 然后使用下面公式更新w, 使得$f_w(x)$尽可能逼近$\hat{y}$:

$$w^+ = w - \eta \frac{\partial E}{\partial w}$$

> 真实值: 令 $x_1 = 1$ 　$x_2 = 0.5$
>
> w_1 w_2 w_3 w_4 w_5 w_6 真实值为 1, 2, 3, 4, 0.5 0.6
>
> y计算真实值可得4

模拟Back Propagation流程:

已知 $x_1 = 1$, $x_2 = 0.5$, $\hat{y} = 4$ (label), w_1 w_2 w_3 w_4 w_5 未知,

随机初始化 $w_1 = 0.5$, $w_2 = 1.5$, $w_3 = 2.3$, $w_4 = 3$, $w_5 = 1$, $w_6 = 1$

① 前馈运算:

计算 h_1, h_2, y和误差E:

$h_1 = w_1 x_1 + w_2 x_2 = 1.25$

$h_2 = w_3 x_1 + w_4 x_2 = 3.8$

$y = w_5 h_1 + w_6 h_2 = 5.05$

$E = \frac{1}{2}(y - \hat{y})^2 = 0.55125$

② 反向传播

$$\frac{\partial E}{\partial w_5}=\frac{\partial E}{\partial y}\frac{\partial y}{\partial w_5}=\left[2\times\frac{1}{2}(y-\hat{y})\times 1\right]\times\left[h_1\right]=(5.05-4)\times 1.25=1.3125$$

$$w_5^{+}=w_5-\eta\frac{\partial E}{\partial w_5}$$

$$=1-0.1\times 1.3125=0.86875$$

同理可得 $\frac{\partial E}{\partial w_6}=\frac{\partial E}{\partial y}\frac{\partial y}{\partial w_6}=1.05\times 3.8=3.99$

$$w_6^{+}=w_6-\eta\frac{\partial E}{\partial w_6}$$

$$=1-0.1\times 3.99=0.601$$

$$\frac{\partial E}{\partial w_1}=\frac{\partial E}{\partial y}\cdot\frac{\partial y}{\partial h_1}\cdot\frac{\partial h_1}{\partial w_1}$$

使用未更新之前值 w_5 而不是 w_5^{+}

$$=1.05\times w_5\times x_1=1.05\times 1\times 1=1.05$$

$$w_1^{+}=w_1-\eta\frac{\partial E}{\partial w_1}$$

$$=0.5-0.1\times 1.05=0.395$$

同理可得 $w_2^{+}=1.4475$　$w_3^{+}=2.195$　$w_4^{+}=2.9475$

③ 前馈运算 2.

$$h_1=w_1^{+}x_1+w_2^{+}x_2=1.11875$$

$$h_2=w_3^{+}x_1+w_4^{+}x_2=3.66875$$

$$y=w_5^{+}h_1+w_6^{+}h_2=3.1768$$

$$E=\frac{1}{2}(y-\hat{y})^2=\frac{1}{2}(3.1768-4)^2=0.3388$$

这次误差比上一次前馈误差0.55125小.