Sol^n : given examples

a	b	d
1	0	1
0	1	0

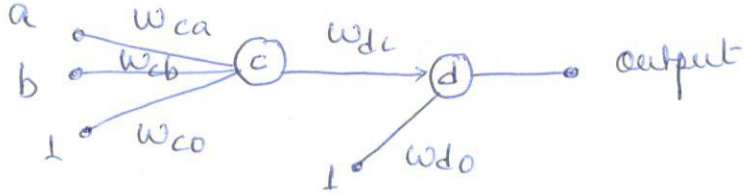
learning rate $\eta = 0.3$

momentum $\alpha = 0.9$

Initial values of weights

$w_{ca} = w_{cb} = w_{co} = w_{dc} = w_{do}$

$= 0.1$



$$net_1 = \sum_{i=0}^{n} w_i x_i$$

Forward phase

$$= w_0 x_0 + w_1 x_1 + w x_2$$

considering
$a=1, \; b=0 \longrightarrow$

$$= w_{co} + w_{ca} a + w_{cb} b = 0.2$$

$$\sigma(net_1) = \frac{1}{1+\exp^{-net_1}} = 0.5498$$

$$net_2 = w_{dc} c + w_{do} = 0.1550$$

$$\sigma(net_2) = \frac{1}{1+\exp^{-net_2}} = 0.5387$$

Backward phase

weight update for output unit weights (without momentum)

$$\Delta w_{ji} = \eta \underbrace{(d_i - O_i)\, O_i (1 - O_i)}_{\delta_j} x_{ji}$$

$$\Delta w_{ji} = \eta \delta_j x_{ji}$$

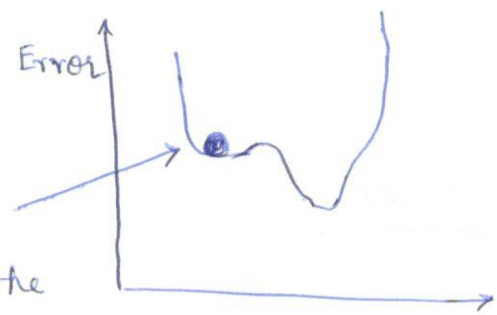
weight update with momentum

$(n-1)^{th}$ iteration

$$\Delta w_{ji}(n) = \eta \delta_j x_{ji} + \alpha \Delta w_{ji}(n-1)$$

$0 \leq \alpha < 1$: constant : momentum

* The momentum helps gradually increasing the step size of search in regions where gradient is and in turn speeding the convergence.



analogy to a ball rolling down the error surface. Momentum helps to keep the ball rolling through flat surfaces (local minima). Without momentum ball would stop in such

---

$$\Delta w_{dc} = 0.3 \times 1 (1 - 0.5387)(1 - 0.5387) \times 0.5498$$

$$= 0.0351$$

$$\Delta w_{d0} = 0.3 \times 1 (1 - 0.5387)(1 - 0.5387) \times 1$$

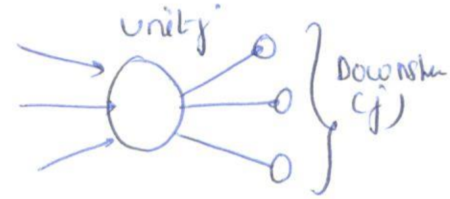
$$= 0.0638$$

with momentum

weight update for hidden unit

$$\Delta w_{ji} = \eta \delta_j x_{ji}$$

$$\delta_j = o_j(1-o_j) \sum_{k \in Downstream(j)} \delta_k w_{kj}$$



Downstream (j) : set o units whose immediate inputs include the output of unit j

$$\delta_k = o_j(1-o_j)(d_j - o_j)$$

as k is the output node.

$$\delta_c = 0.5498(1-0.5498) \times$$

$$[\delta_d w_{dc} + \delta_d w_{do}]$$

$$= 0.00011$$

$$\Delta w_{ca} = 0.3 \times 0.00011 \times 1 = 0.00003$$

$$\Delta w_{cb} = 0.3 \times 0.00011 \times 0 = 0$$

$$\Delta w_{co} = 0.3 \times 0.00011 \times 1 = 0.00003$$

similarly weights with momentum can be determined.

---

finally update the weights.

$$w_{ji} \leftarrow w_{ji} + \Delta w_{ji}$$

$$\Delta w_{dc} = 0.0351 + 0.1 = 0.1351$$

$$\Delta w_{do} = 0.1 + 0.0638 = 0.1638$$

$$\Delta w_{ca} = 0.1 + 0.00003 = 0.1000$$