MLP with energy based gradient descent

Definition:

x_k -input

E_j - energy of the j^{th} node

$\gamma_{ij} = \pm 1$, for representing excitatory and inhibitory connection

c_1, c_2, c_3= constants

w_{jk} - weights connecting input to hidden layer

w_{ij}- output weights connecting the hidden neurons to output layer

b_j^f - bias of hidden layer

b_i^s- bias of the output layer

h_j^f- net input to hidden neuron layer

$V_j = g(h_j^f)$ - output of hidden neuron layer

h_i^s - net input to output layer

$y_i = g(h_j^s)$- output of output layer

d = desired output

g - sigmoid function

Relation of the weights and bias to energy

** A simplification assumed here is that from the leaf node a fraction of the total energy (E) goes for maintaining both the weights (w_{jk},w_{ij})and the bias(b_j^f). This fraction (c_1, c_2, c_3 in this case) are constants. Only the weights connecting the vascular leaf nodes upwards toward the root node will be updated. (During the discussion, we had originally thought of updating these fractions also.)

$$w_{jk} = \gamma_{jk} c_1 E_j \tag{1}$$

$$w_{ij} = \gamma_{ij} c_2 E_j \tag{2}$$

$$b_j^f = 1 - c_3 E_j \tag{3}$$

$$h_j^f = \sum_k w_{jk} x_k - b_j^f \tag{4}$$
$$= c_1 \sum_k \gamma_{jk} E_j - (1 - c_3 E_j) \tag{5}$$
$$= c_3 E_j + c_1 \sum_k \gamma_{jk} E_j - 1 \tag{6}$$

$$h_i^s = \sum_j w_{ij} V_j - b_i^s \tag{7}$$
$$= c_2 \sum_j \gamma_{ij} E_j V_j - b_i^s \tag{8}$$

The cost function E_p is given by

$$Ep = \frac{1}{2} \sum_i \|d_i - y_i\|^2 + \sum_j |E_j| \tag{9}$$

$$\Delta E_j = -\eta \frac{\partial E_p}{\partial E_j} \tag{10}$$

$$e_i = d_i - y_i \tag{11}$$

$$\frac{\partial E_p}{\partial E_j} = \sum_i e_i (-1) \frac{\partial y_i}{\partial E_j} + \frac{E_j}{|E_j|} \tag{12}$$

$$\frac{\partial y_i}{\partial E_j} = y_i (1 - y_i) \frac{\partial h_i^s}{\partial E_j} \tag{13}$$

$$\frac{\partial h_i^s}{\partial E_j} = c_2 \left(\gamma_{ij} V_j + \gamma_{ij} E_j V_j (1 - V_j) \frac{\partial h_j^f}{\partial E_j} \right) \tag{14}$$

$$\frac{\partial h_j^f}{\partial E_j} = c_3 + c_1 \sum_k \gamma_{jk} \tag{15}$$

Combining equations 11 to 15,

$$\Delta E_j = \eta \sum_i e_i y_i (1 - y_i) c_2 \left(\gamma_{ij} V_j + \gamma_{ij} E_j V_j (1 - V_j)(c_3 + c_1 \sum_k \gamma_{jk}) \right) - \eta \frac{E_j}{|E_j|} \tag{16}$$
$$= \eta \sum_i e_i y_i (1 - y_i) c_2 \left(\gamma_{ij} V_j \left(1 + E_j (1 - V_j) \big(c_3 + c_1 \sum_k \gamma_{jk} \big) \right) \right) - \eta \frac{E_j}{|E_j|} \tag{17}$$

The bias of the second layer, b_i^s

$$\Delta b_i^s = -\eta \frac{\partial E_p}{\partial b_i^s} \tag{18}$$

$$= -\eta e_i y_i (1 - y_i)(-1)(-1) \tag{19}$$

$$= -\eta e_i y_i (1 - y_i) \tag{20}$$

The update equation for energy of leaf nodes and bias of second layer becomes:

$$E_j(n+1) = E_j(n) + \Delta E_j \tag{21}$$

$$b_i^s(n+1) = b_i^s(n) + \Delta b_i^s \tag{22}$$