ENEE434 Spring 2005
Homework3 Solution: Y.J. 02/24/05

Problem 1: (E11.6 , pp.11-45 of the textbook.)

i) Find the square error $(e)^2$:

Output from first layer of neurons:

$$a^1 = f^1\left(W^1 \cdot p + b^1\right) = \log sig\left[\begin{bmatrix} w_{1,1}^1 \\ w_{2,1}^1 \end{bmatrix} \cdot [p] + \begin{bmatrix} b_1^1 \\ b_2^1 \end{bmatrix}\right] = \begin{bmatrix} \log sig\left(w_{1,1}^1 \cdot p + b_1^1\right) \\ \log sig\left(w_{2,1}^1 \cdot p + b_2^1\right) \end{bmatrix}$$

Output from the second layer of neuron:

$$a^2 = f^2\left(W^2 \cdot a^1 + b^2\right) = f^2\left\{\begin{bmatrix} w_{1,1}^2 & w_{1,2}^2 \end{bmatrix} \cdot \begin{bmatrix} a_1^1 \\ a_2^1 \end{bmatrix} + b^2\right\} = \begin{bmatrix} w_{1,1}^2 & w_{1,2}^2 \end{bmatrix} \cdot \begin{bmatrix} \log sig\left(w_{1,1}^1 \cdot p + b_1^1\right) \\ \log sig\left(w_{2,1}^1 \cdot p + b_2^1\right) \end{bmatrix} + b^2$$

Output of the neural network:

$$a = a^2 = w_{1,1}^2 \cdot \log sig\left(w_{1,1}^1 \cdot p + b_1^1\right) + w_{1,2}^2 \cdot \log sig\left(w_{2,1}^1 \cdot p + b_2^1\right) + b^2$$

Target of the neural network: $t = 1 + \sin\left(\frac{\pi}{4} \cdot p\right)$

Square error:

$$e^2 = (t-a)^2 = \left\{\left[1 + \sin\left(\frac{\pi}{4} \cdot p\right)\right] - \left[w_{1,1}^2 \cdot \log sig\left(w_{1,1}^1 \cdot p + b_1^1\right) + w_{1,2}^2 \cdot \log sig\left(w_{2,1}^1 \cdot p + b_2^1\right) + b^2\right]\right\}^2$$

with $p = 1$,

$$e^2 = (t-a)^2 = \left(1 + \sin\frac{\pi}{4} - w_{1,1}^2 \cdot \log sig\left(w_{1,1}^1 + b_1^1\right) - w_{1,2}^2 \cdot \log sig\left(w_{2,1}^1 + b_2^1\right) - b^2\right)^2, \text{ or}$$

$$e^2 = (t-a)^2 = \left(1 + \sin\frac{\pi}{4} - \frac{w_{1,1}^2}{1 + e^{-\left(w_{1,1}^1 + b_1^1\right)}} - \frac{w_{1,2}^2}{1 + e^{-\left(w_{2,1}^1 + b_2^1\right)}} - b^2\right)^2$$

ii) Find $\frac{\partial\left(e^2\right)}{\partial w_{1,1}^1}$:

With initial conditions:

$W^1(0) = \begin{bmatrix} -0.27 \\ -0.41 \end{bmatrix}$, $b^1(0) = \begin{bmatrix} -0.48 \\ -0.13 \end{bmatrix}$, $W^2(0) = \begin{bmatrix} 0.09 & -0.17 \end{bmatrix}$, $b^2(0) = [0.48]$, and $a(0) = [0.446]$.

Target t: $t = 1 + \sin\left(\frac{\pi}{4} \cdot p\right) = 1.707$

$$\frac{\partial(e^2)}{\partial w_{1,1}^1} = -2 \cdot (t-a) \cdot \frac{\partial a}{\partial w_{1,1}^1} = -2.522 \cdot \frac{\partial a}{\partial w_{1,1}^1}$$

$$\frac{\partial a}{\partial w_{1,1}^1} = w_{1,1}^2 \cdot \frac{\partial\left[\log sig\left(w_{1,1}^1 \cdot p + b_1^1\right)\right]}{\partial w_{1,1}^1} = 0.09 \cdot \frac{\partial\left[\log sig\left(w_{1,1}^1 \cdot p + b_1^1\right)\right]}{\partial w_{1,1}^1}$$

$$\frac{\partial\left[\log sig\left(w_{1,1}^1 \cdot p + b_1^1\right)\right]}{\partial w_{1,1}^1} = \frac{\exp\left(-w_{1,1}^1 \cdot p - b_1^1\right)}{\left[1+\exp\left(-w_{1,1}^1 \cdot p - b_1^1\right)\right]^2} \cdot p = \frac{\exp(0.75)}{\left[1+\exp(0.75)\right]^2} \cdot 1 = 0.21789$$

$$\frac{\partial(e^2)}{\partial w_{1,1}^1} = -2.522 \times 0.09 \times 0.21789 = -0.0495$$

iii) From the textbook: pp.11-16~17:

$s^1 = \begin{bmatrix} -0.0495 \\ 0.0997 \end{bmatrix}$ and $s_{1,1}^1 = -0.0495$

$w_{1,1}^1(1) = w_{1,1}^1(0) - \alpha \cdot s_{1,1}^1$

also, from the textbook: (pp.11-9) $w_{1,1}^1(1) = w_{1,1}^1(0) - \alpha \cdot \dfrac{\partial(e^2)}{\partial w_{1,1}^1}$

clearly, $\dfrac{\partial(e^2)}{\partial w_{1,1}^1} = s_{1,1}^1 = -0.0495$, the two equations to find the next set of weights are the same.

Problem 2: (E11.10 , pp.11-47 of the textbook.)

$a^1 = f^1\left(w^1 \cdot p + b^1\right)$ and $a^2 = f^2\left(w^2 \cdot a^1 + w^{2,1} \cdot p + b^2\right)$

w^1, b^1, w^2, and b^2 are standard weighs and biases, therefore, the updated equation do not change, as shown in equation 11.27 and 11.28 from the textbook (pp.11-10).

The additional equation for $w^{2,1}$ needs to be derived.

$$\frac{\partial \hat{F}}{\partial w^{2,1}} = \frac{\partial \hat{F}}{\partial n^2} \cdot \frac{\partial n^2}{\partial w^{2,1}} = s^2 \cdot p$$

Therefore, the updated equations are:

$W^m(k+1) = W^m(k) - \alpha \cdot s^m\left(a^{m-1}\right)$ with m=1, 2

$b^m(k+1) = b^m(k) - \alpha \cdot s^m$ with m=1, 2

$w^{2,1}(k+1) = w^{2,1}(k) - \alpha \cdot s^2 \cdot p$