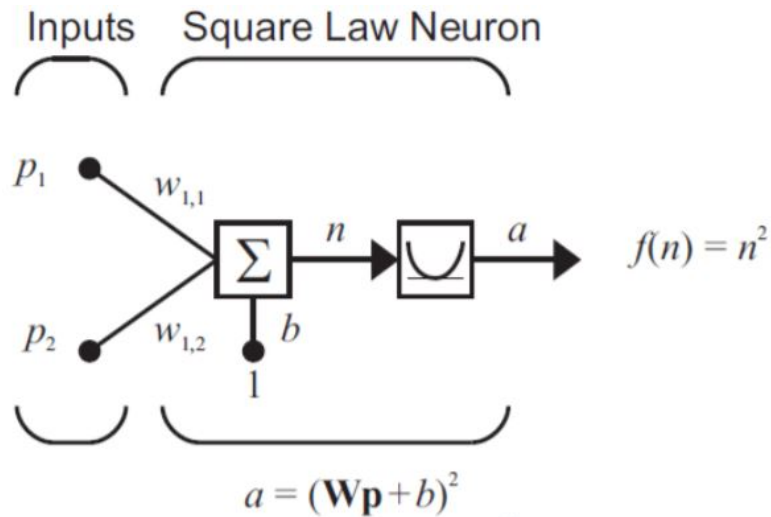


Κεφάλαιο 2

Πρόβλημα 2

Στο πρόβλημα αυτό εκπαιδεύουμε το δίκτυο της εικόνας 2.1 χρησιμοποιώντας τον Standard Backpropagation αλγόριθμο (approximate steepest descent), με αρχικά βάρη και bias : $W(0) = \begin{bmatrix} 1 & -1 \end{bmatrix}$, $b(0) = 1$, δοθέντος ενός ζευγαριού εισόδου/στόχου $\{p = \begin{bmatrix} 1 & 1 \end{bmatrix}^T, t = \begin{bmatrix} 0 \end{bmatrix}\}$. Τα δεδομένα εξόδου από την επίλυση του προβλήματος και της εκτέλεσης μίας επανάληψης του αλγορίθμου απεικονίζονται συγκεντρωτικά στον πίνακα 2.1.



Εικόνα 2.1: Δίκτυο προβλήματος 2

Αρχικά, υπολογίζουμε την παράγωγο της συνάρτησης μεταφοράς :

1.

$$f'(n) = 2n \quad (2.1)$$

2.1 Propagating the input forward through the network

1.

$$n = \mathbf{W}\mathbf{p} + b = W_{1,1}p_1 + W_{1,2}p_2 + b = 1 \cdot 1 + (-1) \cdot 1 + 1 = 1 \quad (2.2)$$

2.

$$a^0 = p = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (2.3)$$

Πίνακας 2.1: Δεδομένα Εξόδου προβλήματος 2

Μεταβλητή	Τιμή		
n	1		
a^0		1	
		-1	
a	1		
e	-1		
s	4		
∇e^2		2	
		2	
$W(1)$	0.6	-1.4	
$b(1)$	0.6		

3.

$$a = (Wp + b)^2 = f(Wp + b) = f(W_{1,1}p_1 + W_{1,2}p_2 + b) = f(n) = f(1) = 1^2 = 1 \quad (2.4)$$

2.2 Computing error

$$e = t - a = 0 - 1 = -1 \quad (2.5)$$

2.3 Propagating the sensitivities backward through the network

$$s = -2f'(n)(t - a) = -2 \cdot 2 \cdot 1 \cdot (-1) = 4 \quad (2.6)$$

2.4 Computing the gradient of the squared error with respect to the weights and bias

$$\nabla e^2 = -2e \cdot p = -2 \cdot (-1) \cdot \begin{vmatrix} 1 \\ 1 \end{vmatrix} = \begin{vmatrix} 2 \\ 2 \end{vmatrix} \quad (2.7)$$

2.5 Updating the weights and bias (assume a learning rate of $\alpha = 0.1$)

1.

$$W(1) = W(0) - \alpha s (a^0)^T = \begin{vmatrix} 1 & -1 \end{vmatrix} - 0.1 \cdot 4 \cdot \begin{vmatrix} 1 & 1 \end{vmatrix} = \begin{vmatrix} 0.6 & -1.4 \end{vmatrix} \quad (2.8)$$

2.

$$b(1) = b(0) - \alpha \cdot s = 1 - 0.1 \cdot 4 = 0.6 \quad (2.9)$$