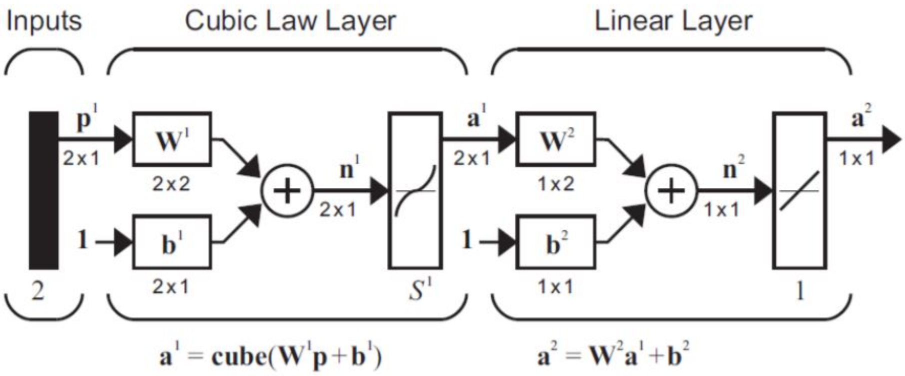
Κεφάλαιο 6

Πρόβλημα 6

Στο πρόβλημα αυτό ζητείται η εκτέλεση μίας επανάληψης του standard steepest descent backpropagation αλγορίθμου, χρησιμοποιώντας πράξεις πινάκων, για το πολυεπίπεδο αντίληπτρο δίκτυο που απεικονίζεται στην εικόνα 6.1 , με $a = 0.5$ και αρχικά βάρη και bias : $w^1(0) = \begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix}, b^1(0) = \begin{vmatrix} 1 \\ 2 \end{vmatrix}, w^2(0) = \begin{vmatrix} 1 & 1 \end{vmatrix}, b^2(0) = \begin{vmatrix} 1 \end{vmatrix}$, δοθέντος ενός ζευγαριού εισόδου/στόχου $\{p = \begin{vmatrix} -1 \\ 1 \end{vmatrix}, t = \begin{vmatrix} -1 \end{vmatrix}\}$. Η συνάρτηση μεταφοράς του κρυφού επιπέδου είναι $f(n) = n^3$. Τα δεδομένα εξόδου από την επίλυση του προβλήματος και της εκτέλεσης μίας επανάληψης του αλγορίθμου απεικονίζονται συγκεντρωτικά στον πίνακα 6.1 .



Εικόνα 6.1: *Δίκτυο προβλήματος 6*

1. Επισημαίνουμε πως το n στο 1^o επίπεδο είναι δισδιάστατο .

$$f^1(n) = n^3 \tag{6.1}$$

καθώς

$$a^1 = cube(W^1 p + b^1) \tag{6.2}$$

2.

$$f^2(n) = n \tag{6.3}$$

καθώς

$$a^2 = W^2 a^1 + b^2 \tag{6.4}$$

Αρχικά, υπολογίζουμε τις παραγώγους των συναρτήσεων μεταφοράς :

1.
$$f'^1(n^1) = 3(n^1)^2 \tag{6.5}$$

2.
$$f'^2(n^2) = 1 \tag{6.6}$$

Έπειτα διαβιβάζουμε την είσοδο μέσω του δικτύου :

1.
$$n^1 = W^1 p + b^1 = \begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} \cdot \begin{vmatrix} -1 \\ 1 \end{vmatrix} + \begin{vmatrix} 1 \\ 2 \end{vmatrix} = \begin{vmatrix} -2 \\ -1 \end{vmatrix} + \begin{vmatrix} 1 \\ 2 \end{vmatrix} = \begin{vmatrix} -1 \\ 1 \end{vmatrix} \tag{6.7}$$

2.
$$a^0 = p = \begin{vmatrix} -1 \\ 1 \end{vmatrix} \tag{6.8}$$

3.
$$a^1 = f^1(W^1 p + b^1) = f^1(n^1) = f^1(\begin{vmatrix} -1 \\ 1 \end{vmatrix}) = \begin{vmatrix} -1 \\ 1 \end{vmatrix}^3 = \begin{vmatrix} -1 \\ 1 \end{vmatrix} \tag{6.9}$$

4.
$$n^2 = W^2 a^1 + b^2 = \begin{vmatrix} 1 & 1 \end{vmatrix} \cdot \begin{vmatrix} -1 \\ 1 \end{vmatrix} + \begin{vmatrix} 1 \end{vmatrix} = 1 \tag{6.10}$$

5.
$$a^2 = f^2(W^2 a^1 + b^2) = f^2(n^2) = f^2(1) = 1 \tag{6.11}$$

6.
$$e = t - a^2 = \begin{vmatrix} -1 \end{vmatrix} - \begin{vmatrix} 1 \end{vmatrix} = -1 - 1 = -2 \tag{6.12}$$

Επιπροσθέτως, υπολογίζουμε τις ευαισθησίες του δικτύου :

1.
$$s^2 = -2f'^2(n^2)e = -2 \cdot 1 \cdot -2 = 4 \tag{6.13}$$

2.
$$s^1 = f'^1(n^1)(W^2)^T s^2 = 3 \cdot \begin{vmatrix} -1 \\ 1 \end{vmatrix}^2 \cdot \begin{vmatrix} 1 \\ 1 \end{vmatrix} \cdot 4 = \begin{vmatrix} 12 \\ 12 \end{vmatrix} \tag{6.14}$$

Τέλος ενημερώνουμε τα βάρη και την πόλωση (bias) :

1.
$$W^2(1) = W^2(0) - a \cdot s^2 (a^1)^T = \begin{vmatrix} 1 & 1 \end{vmatrix} - 0.5 \cdot 4 \cdot \begin{vmatrix} -1 & 1 \end{vmatrix} = \begin{vmatrix} 3 & -1 \end{vmatrix} \tag{6.15}$$

2.
$$W^1(1) = W^1(0) - a \cdot s^1 (a^0)^T = \begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} - 0.5 \cdot \begin{vmatrix} 12 \\ 12 \end{vmatrix} \cdot \begin{vmatrix} -1 & 1 \end{vmatrix} = \begin{vmatrix} 7 & -7 \\ 7 & -6 \end{vmatrix} \tag{6.16}$$

Πίνακας 6.1: *Δεδομένα Εξόδου προβλήματος 6*

Μεταβλητή	**Τιμή**
n^1	$\begin{vmatrix} -1 \\ 1 \end{vmatrix}$
n^2	1
a^0	$\begin{vmatrix} -1 \\ 1 \end{vmatrix}$
a^1	$\begin{vmatrix} -1 \\ 1 \end{vmatrix}$
a^2	1
s^1	$\begin{vmatrix} -6 \\ -6 \end{vmatrix}$
s^2	4
e	-2
$W^1(1)$	$\begin{vmatrix} 7 & -7 \\ 7 & -6 \end{vmatrix}$
$W^2(1)$	$\begin{vmatrix} 3 & -1 \end{vmatrix}$
$b^1(1)$	$\begin{vmatrix} -5 \\ -4 \end{vmatrix}$
$b^2(1)$	-1

3.

$$b^2(1) = b^2(0) - a \cdot s^2 = \begin{vmatrix} 1 \end{vmatrix} - 0.5 \cdot 4 = -1 \tag{6.17}$$

4.

$$b^1(1) = b^1(0) - a \cdot s^1 = \begin{vmatrix} 1 \\ 2 \end{vmatrix} - 0.5 \cdot \begin{vmatrix} 12 \\ 12 \end{vmatrix} = \begin{vmatrix} -5 \\ -4 \end{vmatrix} \tag{6.18}$$