

Ασκήσεις Φροντιστηρίου «Υπολογιστική Νοημοσύνη Ι»

4^ο Φροντιστήριο

Πρόβλημα 1^ο

Ο πίνακας συσχέτισης \mathbf{R}_x του διανύσματος εισόδου $\mathbf{x}(n)$ στον LMS αλγόριθμο

$$\mathbf{R}_x = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$

ορίζεται ως:

Ορίστε το διάστημα των τιμών της παραμέτρου μάθησης η του LMS αλγόριθμου έτσι ώστε ο αλγόριθμος να συγκλίνει

α) στη μέση τιμή, $0 < \eta < 2 / \lambda_{\max}$

β) στη μέση τετραγωνική τιμή, $0 < \eta < 2 / \text{tr}[\mathbf{R}_x]$ ή $0 < \eta < 2 / \|\mathbf{x}(n)\|_2$

Λύση

Βρίσκουμε τις ιδιοτιμές του \mathbf{R}_x ως εξής:

$$\det(\mathbf{R}_x - \lambda \mathbf{I}) = \begin{vmatrix} 1 - \lambda & 0,5 \\ 0,5 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^2 - 0,5^2 = 0,75 - 2\lambda + \lambda^2 = 0$$

Οπότε έχουμε :

$$\lambda = 1/2, 3/2$$

α)

Ο αλγόριθμος LMS συγκλίνει στην μέση τιμή αν η παράμετρος μάθησης ικανοποιεί την συνθήκη :

$$0 < \eta < 2/\lambda_{\max}$$

Στην συγκεκριμένη περίπτωση έχουμε $\lambda_{\max}=3/2$ οπότε :

$$0 < \eta < 4/3$$

β)

Ο αλγόριθμος LMS συγκλίνει στη μέση τετραγωνική τιμή αν :

$$0 < \eta < \frac{2}{\text{tr}[\mathbf{R}_x]}$$

Στην συγκεκριμένη περίπτωση έχουμε $\text{tr}[\mathbf{R}_x] = \lambda_{\max} + \lambda_{\min} = 3/2 + 1/2 = 2$

Οπότε:

$$0 < \eta < 1$$

Πρόβλημα 2^ο

Ο κανονικοποιημένος LMS αλγόριθμος περιγράφεται από την παρακάτω αναδρομική εξίσωση των βαρών:

$$\hat{w}(n+1) = \hat{w}(n) + \frac{\eta}{\|x(n)\|^2} e(n)x(n),$$

όπου η είναι μία θετική σταθερά και $\|x(n)\|$ είναι η Ευκλείδεια νόρμα του διανύσματος εισόδου $x(n)$. Το σήμα λάθους δίνεται από τη σχέση

$$e(n) = d(n) - w^T(n)x(n),$$

όπου $d(n)$ είναι η επιθυμητή απόκριση. Δείξτε ότι για να συγκλίνει ο κανονικοποιημένος LMS στη μέση τετραγωνική τιμή πρέπει να ισχύει

$$0 < \eta < 2$$

Λύση

Στον απλό LMS έχουμε ότι :

$$w(n+1) = w(n) + \eta e(n)x(n) \quad (1)$$

οπότε η σύγκλιση στο μέσο τετράγωνο γίνεται όταν :

$$0 < \eta < \frac{2}{\|x(n)\|^2} \quad (2)$$

Στον κανονικοποιημένο LMS έχουμε :

$$w(n+1) = w(n) + \eta \frac{e(n)x(n)}{\|x(n)\|^2} \quad (3)$$

Συγκρίνοντας την (1) και την (3) μπορούμε να γράψουμε :

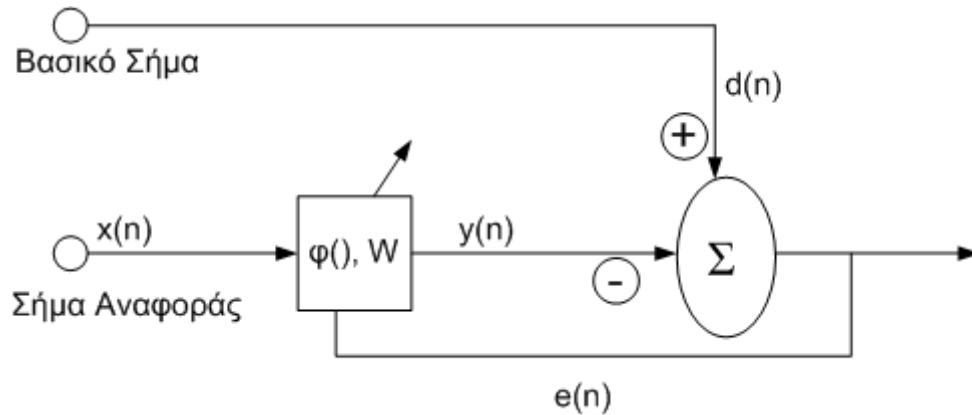
$$\eta' = \eta \|x(n)\|^2 \quad (4)$$

Χρησιμοποιώντας την (2) και την (4) έχουμε:

$$0 < \eta < \frac{2}{\|x(n)\|^2} \Leftrightarrow 0 < \frac{\eta'}{\|x(n)\|^2} < \frac{2}{\|x(n)\|^2} \Leftrightarrow 0 < \eta' < 2$$

Πρόβλημα 3^ο

Ο LMS αλγόριθμος χρησιμοποιείται για να υλοποιηθεί ο δύο εισόδων, με ένα βάρος, προσαρμοστικός εξουδετερωτής θορύβου που φαίνεται στο παρακάτω σχήμα. Καθορίστε τις εξισώσεις που περιγράφουν τη λειτουργία αυτού του αλγόριθμου.

**Λύση**

Το βασικό σήμα $d(n)$ είναι το επιθυμητό σήμα. Το σήμα αναφοράς $x(n)$ είναι το σήμα εισόδου. Το σήμα $e(n)$ είναι το σήμα λάθους : $e(n) = d(n) - y(n)$

Όπου $y(n) = \varphi(w(n)x(n))$.

Η $\varphi(\cdot)$ είναι η συνάρτηση ενεργοποίησης του νευρωνικού δικτύου , οπότε ισχύει :

$$\frac{\partial e(n)}{\partial w(n)} = \frac{\partial}{\partial w(n)} (d(n) - \varphi(w(n)x(n))) = -\varphi'(w(n)x(n)) \cdot x(n)$$

Για την ανανέωση των βαρών έχουμε :

$$w(n+1) = w(n) - \eta e(n) \frac{\partial e(n)}{\partial w(n)} = w(n) + \eta e(n) \varphi'(w(n)x(n)) \cdot x(n)$$

Εαν ο νευρώνας είναι γραμμικός, δηλαδή $\varphi'(1) = 1$, τότε έχουμε τον απλό LMS

$$w(n+1) = w(n) + \eta e(n)x(n)$$

Πρόβλημα 4^ο

Ένας γραμμικός προβλεπτής με διάνυσμα εισόδου του να αποτελείται από τα δείγματα $x(n-1), x(n-2), \dots, x(n-p)$ όπου p είναι η τάξη της πρόβλεψης. Πρέπει να χρησιμοποιήσουμε τον LMS αλγόριθμο για να προβλέψουμε την τιμή $\hat{x}(n)$ από τα δείγματα $x(n)$. Γράψτε τις αναδρομικές εξισώσεις που απαιτούνται για τον υπολογισμό των βαρών w_1, w_2, \dots, w_p .

Λύση

Το σήμα εισόδου του προβλεπτή είναι:

$$X^T(n-1) = [x(n-1), x(n-2), \dots, x(n-p)]$$

Η επιθυμητή απόκριση είναι η $\mathbf{d}(\mathbf{n})=\mathbf{x}(\mathbf{n})$

Χρησιμοποιώντας τον LMS έχουμε:

$$\hat{x}(n) = \mathbf{w}^T(n)\mathbf{X}'(n-1)$$

$$e(n) = x(n) - \hat{x}(n)$$

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \eta e(n)\mathbf{X}'(n-1)$$

Πρόβλημα 5^ο

Έστω ότι έχουμε ένα Νευρωνικό Δίκτυο εμπρόσθιας τροφοδότησης με αρχιτεκτονική 1-2-1. Θέλουμε να εκπαιδεύσουμε το δίκτυο, χρησιμοποιώντας αλγόριθμο πίσω διάδοσης του λάθους να προσεγγίζει την ακόλουθη συνάρτηση:

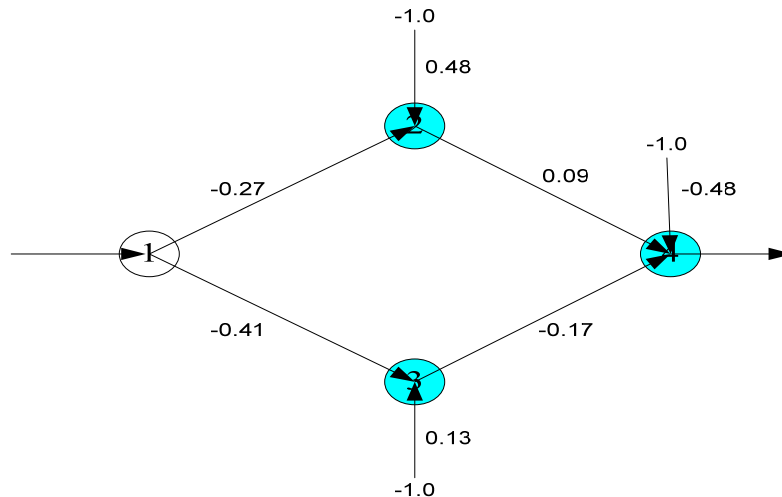
$$g(p) = 1 + \sin\left(\frac{\pi}{4}p\right), \text{ όπου } p \in [-2, 2].$$

Για τους νευρώνες του κρυφού επιπέδου χρησιμοποιούμε σα συνάρτηση ενεργοποίησης τη λογιστική ($\varphi(v) = \frac{1}{1 + e^{-v}}$) ενώ στο επίπεδο εξόδου την γραμμική ($\varphi(v) = v$). Για τις ελεύθερες παραμέτρους του δικτύου θεωρήστε τις ακόλουθες αρχικές τιμές: $w_{21} = -0.27$, $w_{31} = -0.41$, $w_{42} = 0.09$, $w_{43} = -0.17$, $\theta_2 = 0.48$, $\theta_3 = 0.13$ και $\theta_4 = -0.48$.

Χρησιμοποιώντας παράμετρο μάθησης $\eta=0.1$ υλοποιείστε μια επανάληψη του αλγορίθμου Πίσω-Διάδοσης του Λάθους.

Λύση

Το νευρωνικό δίκτυο είναι το εξής:



Η συνάρτηση η οποία θέλουμε να προσεγγίζει το πιο πάνω Νευρωνικό Δίκτυο είναι η

$$g(p) = 1 + \sin\left(\frac{\pi}{4}p\right), \text{ όπου } p \in [-2, 2].$$

Για να μπορέσει το Νευρωνικό Δίκτυο να προσεγγίσει την πιο πάνω συνάρτηση θα πρέπει να εκπαιδευτεί στο διάστημα $[-2, 2]$. Συνήθως επιλέγουμε τιμές εκπαίδευσης του δικτύου ομοιόμορφα στο $[-2, 2]$ εκτός και αν θέλουμε να προσεγγίσουμε μία περιοχή καλύτερα από τις υπόλοιπες οπότε σε αυτή την περιοχή παίρνουμε περισσότερες τιμές εκπαίδευσης απ' ότι στις υπόλοιπες.

Εφόσον μας ζητείται μόνο μια επανάληψη του αλγορίθμου πίσω διάδοσης επιλέγουμε στο $[-2, 2]$ την τιμή 1 για εκπαίδευση (χάριν ευκολίας).

Η έξοδος η οποία θέλουμε να προσεγγίσουμε είναι η:

$$d = g(1) = 1 + \sin\left(\frac{\pi}{4}1\right) \Rightarrow d = 1 + \frac{\sqrt{2}}{2} = 1.707$$

Στη συνέχεια θα πρέπει να βάλουμε ως είσοδο στο νευρωνικό δίκτυο την τιμή 1 και θα υπολογίσουμε την έξοδό του. Μετά θα συγκρίνουμε την έξοδο του νευρωνικού δικτύου με την επιθυμητή έξοδο d και εφαρμόζοντας τις εξισώσεις της πίσω διάδοσης, θα προχωρήσουμε από το επίπεδο εξόδου στο επίπεδο εισόδου ενημερώνοντας (διορθώνοντας) σε κάθε επίπεδο τα βάρη. Στην άσκηση μας ζητείται μόνο μια επανάληψη, διαφορετικά θα συνεχίζαμε την πιο πάνω διαδικασία έως ότου

το νευρωνικό δίκτυο μας δώσει προσέγγιση μικρότερη από το επιθυμητό όριο. Αυτή τη διαδικασία θα πρέπει να την επαναλάβουμε για όλα τα δείγματα εκπαίδευσης.

Εμπρός Πέρασμα:

Για το νευρώνα 2 έχουμε:

$$v_2 = x_1 \cdot w_{21} - \theta_2 = 1 \cdot (-0.27) + (-1) \cdot 0.48 \Rightarrow v_2 = -0.75$$

$$\text{και } y_2 = \varphi(v_2) = \frac{1}{1+e^{-v_2}} = \frac{1}{1+e^{0.75}} = 0.321$$

Για το νευρώνα 3 ισχύει:

$$v_3 = x_1 \cdot w_{31} - \theta_3 = 1 \cdot (-0.41) + (-1) \cdot 0.13 \Rightarrow v_3 = -0.54$$

$$\text{και } y_3 = \varphi(v_3) = \frac{1}{1+e^{-v_3}} = \frac{1}{1+e^{0.57}} = 0.368$$

Για το επίπεδο εξόδου ακολουθούμε την ίδια διαδικασία. Μόνο που σε αυτή την περίπτωση σαν συνάρτηση ενεργοποίησης δε χρησιμοποιούμε την λογιστική απεικόνιση αλλά την ταυτοτική συνάρτηση: $\varphi(v) = v$.

Έτσι για το νευρώνα 4 έχουμε:

$$v_4 = \sum_{i=2}^3 y_i w_{4i} - \theta_4 = 0.09 \cdot 0.32 + (-0.17) \cdot 0.368 + (-1) \cdot (-0.48) \Rightarrow v_4 = 0.446$$

$$\text{και } y_4 = \varphi(v_4) = 0.446.$$

Στη συνέχεια βρίσκουμε το λάθος e στις εξόδους του δικτύου (έχουμε μία έξοδο):

$$e = e_4 = d - y_4 = 1.707 - 0.446 \Rightarrow 1.261$$

Πίσω Πέρασμα:

Υπολογίζουμε τα δ για όλους του υπολογιστικούς νευρώνες του δικτύου, ξεκινώντας από το επίπεδο εξόδου και προχωρώντας προς τα πίσω έως και το πρώτο κρυφό επίπεδο. Στο συγκεκριμένο πρόβλημα έχουμε μόνο ένα νευρώνα στο επίπεδο εξόδου και ένα μόνο κρυφό επίπεδο με δύο κρυφούς νευρώνες: Οπότε, για το νευρώνα 4

έχουμε (θα πρέπει να λάβουμε υπόψη ότι στο νευρώνα εξόδου χρησιμοποιούμε τη γραμμική συνάρτηση ενεργοποίησης):

$$\delta_4(n) = e_4 \cdot \phi'(v_4) = e \cdot 1 \Rightarrow \delta_4(n) = 1.261$$

Εν συνεχεία υπολογίζουμε τις τοπικές κλίσεις για τους δύο νευρώνες του κρυφού επιπέδου (εδώ θα πρέπει να λάβουμε υπόψη ότι οι νευρώνες του κρυφού επιπέδου χρησιμοποιούν τη λογιστική συνάρτηση ενεργοποίησης).

$$\delta_2 = \phi'(v_2) \cdot \delta_4 \cdot w_{42} \Rightarrow \delta_2 = y_2 \cdot (1 - y_2) \cdot \delta_4 \cdot w_{42} \Rightarrow \delta_2 = 0.321 \cdot (1 - 0.321) \cdot 1.261 \cdot 0.09 = 0.0247$$

όμοια υπολογίζουμε και το

$$\delta_3 = -0.0498$$

Στη συνέχεια υπολογίζουμε τις **διορθώσεις στα βάρη και τα κατώφλια** του δικτύου:

Για το νευρώνα 4 έχουμε:

$$\Delta\theta_4 = \Delta w_{40} = n \cdot \delta_4 \cdot y_{40} = 0.1 \cdot 1.261 \cdot (-1) = -0.1261$$

$$\Delta w_{43} = n \cdot \delta_4 \cdot y_3 = 0.1 \cdot 1.261 \cdot 0.368 = 0.0464$$

$$\Delta w_{42} = n \cdot \delta_4 \cdot y_2 = 0.1 \cdot 1.261 \cdot 0.32 = 0.0404$$

Για το νευρώνα 2 έχουμε:

$$\Delta\theta_2 = \Delta w_{20} = n \cdot \delta_2 \cdot y_{20} = -0.00247$$

$$\Delta w_{21} = n \cdot \delta_2 \cdot y_1 = 0.00247$$

Για το νευρώνα 3 έχουμε:

$$\Delta\theta_3 = \Delta w_{30} = n \cdot \delta_3 \cdot y_{30} = 0.00498$$

$$\Delta w_{31} = n \cdot \delta_3 \cdot y_1 = -0.00498.$$

Τέλος υπολογίζουμε τις καινούργιες (διορθωμένες) τιμές των ελεύθερων παραμέτρων σύμφωνα με τον τύπο: $w_{ji}(n+1) = w_{ji}(n) + \Delta w_{ji}(n)$.

Έχουμε:

$$\theta_4(n+1) = \theta_4(n) + \Delta\theta_4(n) = -0.48 - 0.1261 = -0.6061$$

$$w_{42}(n+1) = 0.1304$$

$$w_{43}(n+1) = -0.1236$$

$$\theta_3(n+1) = 0.13498$$

$$w_{31}(n+1) = -0.41498$$

$$\theta_2(n+1) = 0.47753$$

$$w_{21}(n+1) = -0.26753$$

Αυτές είναι και οι νέες τιμές των βαρών και των κατωφλίων του δικτύου.

Σε ένα «πραγματικό» πρόβλημα η παραπάνω διαδικασία θα έπρεπε να συνεχιστεί έως ότου να φτάσουμε σε ένα ϵ μικρότερο από το λάθος προσέγγισης που ζητάμε για όλες τις τιμές που θα χρησιμοποιήσουμε για εκπαίδευση στο διάστημα $[-2,2]$.